

Evaluación Psicológica: Fundamentación general, metodológica y programación docente

Volumen I

Maite Garaigordobil Landazabal

**Evaluación Psicológica: Fundamentación General,
metodológica y programación docente**

Volumen 1

Maite Garaigordobil Landazabal

Proyecto Docente para el Concurso de Acceso para Catedráticos de
Universidad. Area de conocimiento de Personalidad, Evaluación y
Tratamientos Psicológicos. Perfil Evaluación Psicológica.
Facultad de Psicología. Universidad del País Vasco.

22-06-2009

Reservados todos los derechos. El contenido de esta obra está protegido por la Ley, que establece penas de prisión o multas, además de las correspondientes indemnizaciones por daños y perjuicios, para quienes reprodujeren, plagiaren, distribuyeren o comunicaren públicamente, en todo o en parte, una obra literaria, artística o científica, o su transformación, interpretación o ejecución artística fijada en cualquier tipo de soporte o comunicada a través de cualquier medio, sin la preceptiva autorización.

© Maite Garaigordobil Landazabal 2009
Facultad de Psicología.
Avda. Tolosa 70. 20018.
Donostia-San Sebastián

ISBN: 978-84-692-4511-8
Depósito Legal: SS-1014-2009

Impresión:
Zorroaga, S.L.
Avda. Tolosa 70. 20018 Donostia-San Sebastián

PARTE I
FUNDAMENTACIÓN GENERAL Y METODOLÓGICA

INDICE

INTRODUCCIÓN	15
CAPÍTULO 1. EL CONCEPTO DE EVALUACIÓN PSICOLÓGICA	19
1.1. Introducción.....	21
1.2. Precisiones conceptuales: psicodiagnóstico, evaluación y valoración psicológica	22
1.2.1. Tests psicológicos y medición	22
1.2.2. Psicodiagnóstico	24
1.2.3. Evaluación.....	26
1.2.3.1. Evaluación conductual	27
1.2.3.2. Evaluación psicológica.....	28
1.2.4. Valoración	31
1.3. Hacia una pluralidad conceptual en la evaluación psicológica.....	32
1.4. Resumen: El concepto de Evaluación Psicológica	35
CAPÍTULO 2. HISTORIA DE LA EVALUACIÓN PSICOLÓGICA	37
2.1. Introducción.....	39
2.2. Fuentes del Psicodiagnóstico y/o de la Evaluación psicológica: Antecedentes precientíficos	39
2.2.1. Míticos.....	40
2.2.2. Racional-Especulativos	40
2.3. Proceso de constitución del Psicodiagnóstico (S.XIX) y desarrollos del S. XX en la Evaluación Psicológica	41
2.3.1. Estructuración de Anastasi (Anastasi, 1968/1978; Anastasi y Urbina, 1998).....	42
2.3.2. Estructuración de Fernández-Ballesteros (Fernández-Ballesteros, 1992, 1994, 2004).....	44
2.3.3. Estructuración de Pelechano (Pelechano, 1981, 1988)	49
2.4. Notas sobre la crisis y resurgimiento de la Evaluación Psicológica	56
2.4.1. Causas de la crisis del psicodiagnóstico tradicional	56
2.4.2. Indicadores del resurgimiento de la evaluación psicológica.....	57
2.5. Evaluación psicológica: Últimos desarrollos.....	59
2.6. Resumen: Evolución histórica de la evaluación psicológica	61

CAPITULO 3. MODELOS EN EVALUACIÓN PSICOLÓGICA.....	65
3.1. Introducción.....	67
3.2. El Modelo Correlacional, Psicométrico o del Atributo	68
3.3. El Modelo Clínico-Dinámico	72
3.3.1. Perspectiva médico-psiquiátrica	72
3.3.2. Perspectiva psicoanalítica	74
3.4. El Modelo Fenomenológico, Humanista, Existencial, Constructivista.....	79
3.5. El Modelo Conductual.....	81
3.5.1. Enfoque conductual radical: Conductismo de primera generación	85
3.5.2. Enfoque mediacional o de impacto mediatizado: Conductismo de segunda generación	88
3.5.3. Enfoques integradores e interactivos: Conductismo de tercera generación.....	91
3.5.3.1. El modelo Conductual-Cognitivo	92
3.5.3.2. El modelo Conductual Paradigmático.....	95
3.6. El Modelo Cognitivo	98
3.6.1. La evaluación desde el procesamiento de la información.....	102
3.6.2. La evaluación neuropsicológica.....	107
3.7. Resumen: Características diferenciales de los modelos en la evaluación psicológica ..	111
 CAPÍTULO 4. METODOS Y CONTROVERSIAS EN EVALUACIÓN PSICOLÓGICA	 117
4.1. Introducción.....	119
4.2. Principales métodos en la evaluación psicológica.....	119
4.2.1. Método Correlacional	121
4.2.2. Método Experimental.....	123
4.2.3. Método Observacional.....	128
4.3. Alternativas metodológicas polémicas	133
4.3.1. Generalidad - Especificidad de la conducta.....	134
4.3.2. Idiográfico – Nomotético.....	135
4.3.3. Cuantitativo – Cualitativo.....	136
4.3.4. Explicación – Comprensión	137
4.4. Resumen: Principales métodos y alternativas polémicas	138

CAPÍTULO 5. EL PROCESO DE LA EVALUACIÓN PSICOLÓGICA.....	143
5.1. Introducción.....	145
5.2. El proceso en función de las metas de evaluación	145
5.3. Fases y tareas del proceso de evaluación psicológica	147
5.3.1. La perspectiva correlacional: Proceso descriptivo-predictivo	147
5.3.2. La perspectiva experimental: Proceso interventivo-valorativo.....	153
5.3.3. Reflexiones sobre la evaluación psicológica aplicada como proceso correlacional versus experimental	161
5.4. Guías para el proceso de evaluación (GAP).....	164
5.5. Condicionantes que intervienen en el proceso de la evaluación psicológica.....	171
5.5.1. Variables del evaluador	171
5.5.2. Variables del evaluado	172
5.5.3. Variables de la situación.....	174
5.5.4. La calidad del proceso de evaluación psicológica	174
5.5.4.1. La eficacia de los evaluadores en tareas de clasificación, predicción, prescripción de tratamiento.....	174
5.5.4.2. La relevancia de la capacidad de metacognición del evaluador.....	175
5.6. Resumen: Fases, objetivos, tareas y técnicas en el proceso de la evaluación psicológica	178
CAPÍTULO 6. GARANTÍAS CIENTÍFICAS Y ÉTICAS DE LA EVALUACIÓN PSICOLÓGICA	183
6.1. Introducción.....	185
6.2. Garantías científicas de la evaluación psicológica.....	185
6.2.1. Test, técnicas y otros procedimientos de evaluación	185
6.2.2. Procedimiento para la elaboración de un test	185
6.2.3. Interpretación de las puntuaciones utilizadas	188
6.2.4. Propiedades de los tests y de las técnicas estándar de medida	189
6.2.4.1. La fiabilidad	189
6.2.4.2. La validez	193
6.2.5. Teoría de la Generalizabilidad TG	199
6.2.6. Aplicaciones de la Teoría de Respuesta al Ítem TRI	202
6.2.6.1. Los bancos de ítems	204
6.2.6.2. Construcción de tests a medida del sujeto o Test adaptados computerizados	205

6.2.6.3. Tests Referidos al Criterio	206
6.2.6.4. Equiparación de puntuaciones.....	206
6.2.6.5. Sesgo	207
6.3. Garantías éticas en la evaluación psicológica.....	208
6.3.1. El código deontológico del psicólogo.....	208
6.3.2. Directrices asociadas al uso de los tests	214
6.3.2.1. Directrices internacionales para el uso de los tests	214
6.3.2.2. Directrices para la traducción y adaptación de tests.....	218
6.3.2.3. Directrices internacionales para la evaluación informatizada y test aplicados por internet	222
6.4. Resumen: Garantías científicas y éticas en la evaluación psicológica.....	224
 CAPÍTULO 7. TÉCNICAS DE RECOGIDA DE INFORMACIÓN EN EVALUACIÓN PSICOLÓGICA: ENTREVISTA, AUTOINFORMES, OBSERVACIONALES, PROYECTIVAS, SUBJETIVAS, OBJETIVAS, Y PSICOMÉTRICAS.....	
7.1. Introducción.....	233
7.2. La entrevista psicológica.....	237
7.2.1. Definición de la entrevista psicológica	237
7.2.2. Clasificaciones de la entrevista	238
7.2.3. Objetivos de la entrevista	242
7.2.4. La entrevista como instrumento del proceso de evaluación-intervención psicológica	244
7.2.5. Características y habilidades del entrevistador.....	245
7.2.6. Características psicométricas de la entrevista, fuentes de distorsión y estrategias facilitadoras	248
7.3. Los autoinformes	253
7.3.1. Definición y clases de variables evaluadas mediante autoinformes.....	253
7.3.2. Principales tipos de autoinforme.....	254
7.3.3. Fuentes de error de los autoinformes	256
7.4. Las técnicas observacionales	257
7.4.1. Definición y objetivos.....	257
7.4.2. Unidades de observación y de medida	258
7.4.3. Técnicas de registro	260
7.4.4. Consideraciones muestrales	262
7.4.5. Fuentes de error y requisitos científicos de la observación.....	263

7.5. Las técnicas proyectivas	267
7.5.1. Definición e hitos históricos de las técnicas proyectivas	267
7.5.2. Características de las técnicas proyectivas	268
7.5.3. El concepto de proyección.....	269
7.5.4. Clasificación de las técnicas proyectivas	270
7.5.5. Garantías psicométricas y futuro de las técnicas proyectivas.....	272
7.6. Las técnicas subjetivas	275
7.6.1. Definición y características de las técnicas subjetivas.....	275
7.6.2. Técnicas subjetivas representativas	276
7.7. Las técnicas objetivas	279
7.7.1. Definición y características de las técnicas objetivas.....	279
7.7.2. Tipos de técnicas objetivas.....	280
7.7.3. Validez y fiabilidad de las técnicas objetivas	288
7.8. Las técnicas psicométricas	289
7.8.1. Definición y características de los tests psicométricos	289
7.8.2. Clasificación de los tests psicométricos.....	289
7.8.3. Presente y futuro de los tests psicométricos.....	299
7.9. Técnicas de recogida de información en los programas de la asignatura evaluación psicológica en las facultades de psicología españolas.....	301
7.10. Hacia una pluralidad técnico-metodológica en evaluación psicológica	304
7.11. Resumen. Técnicas de recogida de información en evaluación psicológica	306
 CAPÍTULO 8. LINEAS DE FUTURO: EVALUACIÓN INFORMATIZADA Y EVALUACIÓN DE PROGRAMAS	 317
8.1. Introducción.....	319
8.2. La evaluación psicológica informatizada	319
8.2.1. Aplicaciones para la administración de pruebas	321
8.2.1.1. Aplicación informatizada de instrumentos tradicionales	321
8.2.1.2. Aplicación de instrumentos para ser administrados únicamente por ordenador	322
8.2.1.3. Aplicaciones combinando ordenador y vídeo	323
8.2.1.4. Tests adaptativos informatizados (tests a medida).....	323
8.2.1.5. Instrumento auxiliar en la aplicación de técnicas de potencial de aprendizaje	326
8.2.1.6. Desarrollo de programas informáticos integrales	326
8.2.1.7. Aplicación de pruebas basadas en criterios (pruebas de competencia mínima).....	327

8.2.2. Aplicaciones para la corrección - puntuación de pruebas	327
8.2.3. Aplicaciones para la interpretación de pruebas	328
8.2.3.1. Tipos de informes computerizados.....	328
8.2.3.2. Los sistemas expertos	329
8.2.3.3. Programas informáticos en el mercado español.....	331
8.2.4. Impacto y polémica de la evaluación informatizada.....	333
8.2.4.1. La administración informatizada.....	333
8.2.4.2. La interpretación informatizada	336
8.2.4.3. Limitaciones actuales de la evaluación informatizada	337
8.3. La evaluación de programas, intervenciones o tratamientos	339
8.3.1. Bases conceptuales	340
8.3.1.1. Definición y objeto de estudio	340
8.3.1.2. Criterios para la evaluación de programas	342
8.3.1.3. Modelos en la evaluación de programas: evaluación sumativa y formativa.....	343
8.3.1.4. Tareas y fases del proceso en la evaluación de programas	346
8.3.2. Cuestiones metodológicas.....	353
8.3.2.1. Principales diseños en la investigación evaluativa	354
8.3.2.2. Procedimientos e instrumentos para evaluar el cambio	360
8.3.2.3. Problemas metodológicos relacionados con el programa	363
8.3.2.4. El análisis de los datos en la evaluación de los efectos del programa	364
8.3.3. Ámbitos de aplicación de la evaluación de programas en España	365
8.4. Algunas predicciones futuras sobre la evaluación psicológica y el rol del psicólogo en España	367
8.5. Resumen: evaluación informatizada, evaluación de programas y predicciones de futuro.....	369

INDICE DE CUADROS Y TABLAS

Cuadro 1. Relaciones entre medida, evaluación, psicodiagnóstico y valoración en psicología.....	22
Cuadro 2. Evolución histórica de la Evaluación Psicológica	61
Cuadro 3. Organización topográfica de la personalidad	75
Cuadro 4. Características de las Instancias Psíquicas: Ello, Yo y Super-Yo	75
Cuadro 5. Estadios del desarrollo psicoafectivo	76
Cuadro 6. Diferencias entre los Enfoques Conductual y Tradicional de Evaluación	82
Cuadro 7. Complementariedad de la Evaluación Tradicional y la Evaluación Conductual.....	83
Cuadro 8. Características diferenciales de los modelos en la Evaluación Psicológica	112
Cuadro 9. Principales métodos en Evaluación Psicológica	139
Cuadro 10. Alternativas metodológicas polémicas en Evaluación Psicológica	141
Cuadro 11. Proceso de evaluación psicológica: estructura básica, pasos a seguir y guías del proceso de evaluación (GAP)	165
Cuadro 12. Fases y tareas en el proceso de Evaluación Psicológica	179
Cuadro 13. Condicionantes que pueden afectar en el proceso de evaluación.....	181
Cuadro 14. Proceso para la elaboración de un test I	186
Cuadro 15. Proceso para la elaboración de un test II	187
Cuadro 16. Puntuaciones directas, criterio y norma	188
Cuadro 17. Procedimientos de obtención de evidencias de fiabilidad	192
Cuadro 18. Procedimientos de obtención de evidencias de validez	199
Cuadro 19. Propiedades de los tests y de las técnicas estándar de medida	224
Cuadro 20. La teoría de la generalizabilidad	226
Cuadro 21. Aplicaciones de la Teoría de Respuesta al Ítem	226
Cuadro 22. Garantías éticas en la evaluación psicológica.....	227
Cuadro 23. Objetivos de la entrevista inicial.....	244
Cuadro 24. Unidades de observación.....	259
Cuadro 25. Unidades de medida	260
Cuadro 26. Técnicas de registro observacional.....	261
Cuadro 27. Fuentes de error de la observación.....	263
Cuadro 28. Tipos de técnicas proyectivas	271
Cuadro 29. Tests de inteligencia: tests individuales y colectivos.....	291
Cuadro 30. Tests de personalidad: pruebas multirasgo y monorasgo	297

Cuadro 31. Técnicas de recogida de información a incluir en el programa docente de la asignatura Evaluación Psicológica desde la opinión de los docentes	301
Cuadro 32. Test que deben incluirse en el programa docente de la asignatura Evaluación Psicológica desde la opinión de los docentes.....	302
Cuadro 33. Test del listado elaborado por los docentes que forman parte del programa de la asignatura “Evaluación Psicológica” en la Facultad de Psicología de la UPV.....	303
Cuadro 34. Las técnicas en evaluación psicológica y sus taxonomías	306
Cuadro 35. La entrevista psicológica.....	307
Cuadro 36. Los autoinformes	309
Cuadro 37. Las técnicas observacionales	310
Cuadro 38. Las técnicas proyectivas	312
Cuadro 39. Las técnicas subjetivas	313
Cuadro 40. Las técnicas objetivas	313
Cuadro 41. Las técnicas psicométricas	314
Cuadro 42. Características de la Evaluación Formativa y Sumativa.....	343
Cuadro 43. Ventajas e inconvenientes de la evaluación interna o externa	346
Cuadro 44. Tipos de diseños de investigación en evaluación de programas.....	356
Cuadro 45. Paradigmas Cuantitativo-Cualitativo: Características y diseños	356
Cuadro 46. Objetivos de los métodos cuantitativos y cualitativos.....	361
Cuadro 47. Instrumentos de evaluación de programas, intervenciones o tratamientos	361
Cuadro 48. Clasificación de los instrumentos en función de su grado de estandarización	362
Cuadro 49. La evaluación psicológica informatizada	369
Cuadro 50. La evaluación de programas, intervenciones o tratamientos.....	374
Cuadro 51. Algunas predicciones futuras sobre la evaluación psicológica y el rol del psicólogo en España	379

INTRODUCCIÓN

INTRODUCCIÓN

El proyecto docente que se presenta está estructurado en dos partes. La primera parte expone la fundamentación general y metodológica de la asignatura “Evaluación Psicológica” mientras que la segunda parte expone la programación docente y la metodología didáctica de la materia.

La fundamentación general y metodológica se desarrolla en 8 capítulos. El capítulo 1 expone el concepto de la evaluación psicológica y de otros términos afines (medición, psicodiagnóstico, valoración...) con los que en distintos momentos históricos se ha denominado la tarea aplicada del psicólogo evaluador. El capítulo 2 presenta la evolución histórica que ha cursado la evaluación psicológica mencionando algunos de los hitos más relevantes en este decurso. El capítulo 3 pone de relieve la existencia de distintos modelos que coexisten en la evaluación psicológica (psicométrico, clínico-dinámico, fenomenológico, conductual, cognitivo) en relación a los cuales se presentan sus presupuestos teórico-conceptuales, las variables objeto de estudio, los objetivos de la evaluación, los métodos y las técnicas que utilizan, así como sus ámbitos de aplicación. El capítulo 4 describe los principales métodos que se emplean en evaluación psicológica (correlacional, experimental, observacional) planteando una integración de las diversas alternativas polémicas que se han evidenciado en torno a la evaluación (generalidad-especificidad de la conducta, idiográfico-nomotético, cuantitativo-cualitativo, explicación-comprensión).

En el capítulo 5 se identifican las fases del proceso de la evaluación psicológica, tanto cuando se plantea con fines descriptivos-predictivos, como interventivo-valorativos, indicando las tareas que se llevan a cabo en cada fase, e identificando algunos condicionantes que pueden intervenir en el proceso, relacionados con la figura del evaluador, del evaluado o de la situación. El capítulo 6, por un lado, expone cuestiones relacionadas con las garantías científicas de la evaluación psicológica (procedimiento de elaboración de un test, propiedades de los tests y de las técnicas estándar de medida, teoría de la generalizabilidad, teoría de respuesta al ítem...) y, por otro lado, analiza las garantías éticas que debe cumplir la evaluación psicológica (código deontológico, directrices asociadas al uso de los tests...). El capítulo 7 describe las distintas técnicas que se utilizan para recoger información, tales como la entrevista, los autoinformes, las técnicas observacionales, las proyectivas, las subjetivas, las objetivas y las psicométricas. La primera parte del proyecto docente concluye con el capítulo 8 en el que se desarrollan dos grandes líneas de futuro, por un lado, la evaluación psicológica informatizada y, por otro lado, la evaluación de programas, intervenciones o tratamientos. El capítulo concluye perfilando algunas predicciones futuras sobre la evaluación psicológica y el rol del psicólogo en España.

La segunda parte del proyecto presenta la programación docente y la metodología didáctica de la asignatura a través de tres capítulos. El capítulo 9 encuadra el marco legal e institucional del proyecto docente, es decir, por un lado, el marco Europeo de Educación Superior (EEES) y, por otro lado, el marco del nuevo plan de estudios en la Facultad de Psicología de la Universidad del País Vasco. El capítulo 10 se centra en el programa docente de la asignatura “Evaluación Psicológica”, y en este se describen los criterios que se han tenido en cuenta para la elaboración del programa, los objetivos y las competencias que se pretenden desarrollar, la estructura y contenidos del programa, detallando específicamente el programa docente teórico así como las actividades prácticas y trabajos dirigidos que los estudiantes

realizan tanto de forma presencial como no presencial. El capítulo finaliza con la presentación del cronograma de la actividad docente que se desarrollará en las 15 semanas de temporalización de la materia, con la descripción del procedimiento de evaluación de la asignatura, y así mismo se explican las posibilidades de la plataforma virtual “moodle” que se utiliza como apoyo a la docencia presencial de la asignatura. La segunda parte del programa docente concluye con el capítulo 11 en el que se describen diversas fuentes documentales tales como manuales de evaluación psicológica, revistas, bases de datos y direcciones de Internet. Al final se presentan las referencias bibliográficas citadas en el proyecto docente.

CAPÍTULO 1.
EL CONCEPTO DE EVALUACIÓN PSICOLÓGICA

- 1.1. Introducción
- 1.2. Precisiones conceptuales: medición, psicodiagnóstico, evaluación y valoración psicológica
 - 1.2.1. Tests psicológicos y medición
 - 1.2.2. Psicodiagnóstico
 - 1.2.3. Evaluación
 - 1.2.3.1. Evaluación conductual
 - 1.2.3.2. Evaluación psicológica
 - 1.2.4. Valoración
- 1.3. Hacia una pluralidad conceptual en la evaluación psicológica
- 1.4. Resumen: El concepto de "Evaluación y Diagnóstico Psicológicos"

1.1. INTRODUCCIÓN

Como paso previo al estudio sistemático de una materia es condición indispensable definir o, en todo caso, hacer un intento de aproximación a lo que se entiende por Evaluación y Diagnóstico Psicológicos, es decir, delimitar el concepto de la disciplina. La disciplina denominada en los últimos planes de estudio “Evaluación Psicológica” ha tenido diversos enfoques, técnicas, objetivos, aplicaciones e interpretaciones, así como denominaciones. Conocer los cambios que ha ido sufriendo esta disciplina, permitirá situar el actual objeto de su estudio y finalidad, es decir, para qué sirve (Moreno, 2003).

La evaluación psicológica se produce en distintos contextos aplicados consistiendo en una serie de actividades de exploración, medida o análisis de comportamientos o fenómenos psicológicos relativos a un sujeto o grupo específico de sujetos, y esto se realiza mediante un proceso de indagación y toma de decisiones. Así, nada tiene de extraño que tal pluralidad de actividades, contextos y objetivos psicológicos hayan determinado que, a lo largo del tiempo y de los distintos contextos culturales, las acciones evaluativas hayan adoptado distintas denominaciones (Fernández-Ballesteros, 1992a).

¿Qué es la evaluación psicológica? Esta expresión, como traducción del inglés “Psychological Assessment”, es relativamente reciente en Psicología ya que aparece casi exclusivamente en la segunda mitad de nuestro siglo. Su carrera ha sido sin embargo meteórica. Por una parte, su uso se ha extendido extraordinariamente, desplazando a veces a otros que gozaban de indudable “solera”. Por otra parte, su espectro semántico se ha hecho también muy amplio, de tal manera que, a decir de muchos, evaluación psicológica es una expresión comprehensiva que abarcaría dentro de sí lo que se entiende de disciplinas tales como “el psicodiagnóstico”, “el mental o psychological testing”, y “la evaluación conductual”, guardando a su vez estrechas relaciones con la valoración de intervenciones o programas de intervención (*program evaluation*).

La evaluación psicológica es aquella disciplina de la psicología científica que se ocupa de la exploración y análisis del comportamiento (a los niveles de complejidad que se estime oportunos) de un sujeto humano (o grupo) con distintos objetivos (descripción, diagnóstico, selección/predicción, explicación, cambio y/o valoración de los tratamientos o intervenciones aplicadas. La evaluación implica un proceso de solución de problemas y toma de decisiones que comienza con una demanda de un cliente o sujeto (o grupo) a un psicólogo. Este procedimiento implica una serie de actividades científicas y profesionales, realizadas en una secuencia establecida con las que se pretende dar respuesta a esa demanda; entre otras, la administración de una serie de tests, técnicas de medida, y procedimientos de recogida de información (Fernández-Ballesteros, 2004a).

No obstante, el concepto de evaluación psicológica debe ser comprendido en relación a otros términos afines que, en diversos momentos han designado la tarea aplicada del psicólogo evaluador, como son: Psicodiagnóstico, evaluación, valoración, medición... Estos términos son tomados en general como sinónimos, sin embargo, cada uno de ellos tiene connotaciones históricas, conceptuales e incluso metodológicas diferentes por lo que conviene, al menos en un principio, diferenciarlos.

Por ello, y con la finalidad de realizar una clarificación inicial de estos términos, se presenta el Cuadro 1 en el que se definen algunos conceptos relacionados con la tarea evaluadora, tales como: *Medida*, *Psicodiagnóstico*, *Evaluación* y *Valoración* que se

desarrollarán en este capítulo más extensamente. Para facilitar esta clarificación nos remitimos a la diferenciación planteada por Pelechano (1988a).

Cuadro 1. Relaciones entre medida, evaluación, psicodiagnóstico y valoración en psicología (Pelechano, 1988)

CONSTRUCTO	DEFINICIÓN
Medida	Recursos o procedimientos cuantitativos que intentan establecer relaciones entre conceptos abstractos (constructos) e indicadores empíricos. Puede utilizarse una estadística extensiva o intensiva, paramétrica o no paramétrica en función de la parcela objeto de conocimiento
Psicodiagnóstico	Recogida de información procedente de procedimientos no experimentales. Se asienta originariamente en modelos proyectivos y su procedimiento principal de validación es en relación con un criterio previo. Se apela a juicios globales (no analíticos) y no cuantitativos.
Evaluación	Recogida sistemática de información y articulación sistemática de un modelo de funcionamiento personal teórico en función de los datos recogidos. Incluye tanto procedimientos cuantitativos como cualitativos. Insistencia originaria en estudio de caso individual. Incluiría tanto las actividades de medida (tests) como de psicodiagnóstico (pruebas proyectivas), así como de observaciones de campo y entrevistas. El resultado final es una imagen del funcionamiento del individuo.
Valoración	Toma de posición final respecto a una persona, grupo, servicio y/o programa de actuación o técnica terapéutica. Incluye un juicio pronóstico y un plan de actuación con el fin de mejorar y/o reorientar el estado actual de hechos. Posee indudables elementos éticos. Se refiere tanto a aspectos positivos como negativos y, en su uso bibliográfico siempre utiliza procedimientos cuantitativos o modelos susceptibles de cuantificación. Podría y debería incluirse en la parte final de todo informe psicológico al margen del tipo de información y de su modo de recogerla.

Pese a las diferencias semánticas de estos términos, todos ellos mantienen estrechas relaciones entre sí. Además, la evaluación psicológica incluye la medición mediante tests psicológicos, al psicodiagnóstico en el sentido estricto, además de otras técnicas como por ejemplo, la entrevista, la observación-autoobservación, los registros psicofisiológicos, etc. La integración de la información proveniente de estas diversas fuentes conduce a realizar una valoración de la situación actual, así como un pronóstico, y un plan de cambio o recomendaciones, que puede referirse a individuos, grupos, organizaciones, técnicas de intervención, o programas de política social, educativa o sanitaria.

1.2. PRECISIONES CONCEPTUALES: MEDICIÓN, PSICODIAGNÓSTICO, EVALUACIÓN Y VALORACIÓN PSICOLÓGICA

1.2.1. TESTS PSICOLÓGICOS Y MEDICIÓN

La primera utilización del término test mental fue realizada por McKeen Catell en 1890. Para este autor los tests mentales eran unos sistemas normalizados de procedimientos que permitían obtener información objetiva respecto al rendimiento de personas ante la realización de tareas-tipo.

El término test hace referencia a una situación controlada, en la que se recogen muestras de conducta producidas por los sujetos en respuesta a unos estímulos que le son presentados.

Estas respuestas son puntuadas o valoradas según unos criterios, ofreciendo información del lugar que ocupa el sujeto dentro de un grupo de referencia normativo. Aunque con matizaciones, estos elementos se han incluido en diversas definiciones del término test (Anastasi 1968/1978; Anastasi y Urbina, 1998; Cronbach, 1990). Así mismo, Fernández-Ballesteros (1992c; 2004a) define el test como un instrumento sistemático y tipificado que permite la comparación de un sujeto con un grupo normativo.

Pelechano (1988a), a modo de revisión crítica, sintetiza algunas notas características de lo que significan realmente los tests psicológicos, tanto en lo que subyace a ellos, como en lo que se refiere a su futuro:

1. Bajo la denominación tests psicológicos se entiende un proceso de elaboración de pruebas, como la validación, aplicación e interpretación de las mismas. Los tests psicológicos deben de ser utilizados por psicólogos, ya que el uso de otros profesionales no formados en psicología ha afectado al desprestigio actual que aún tienen los tests.
2. El modelo científico básico a la hora de la corrección de los tests y obtención de resultados es el asociacionismo. Así, la puntuación se obtiene con la suma de los aciertos logrados a lo largo de todo el test, en el que cada problema “vale” lo mismo. Un problema importante es que demasiados tests se han construido y se construyen con un empirismo de base careciendo de referencia alguna a constructo teórico o teoría de base, aunque afortunadamente el criterio empirista en la elaboración de pruebas psicológicas está siendo cada vez menos frecuente, y cada vez más, se utilizan construcciones teóricas previas que son las que guían la elaboración de los tests psicológicos.
3. Mientras el modelo de corrección de los tests ha sido asociacionista y mecanicista, el modelo interpretativo y la “versión social” de lo que han significado y significan los tests psicológicos es sustancialista. Las puntuaciones alcanzadas en las pruebas han sido interpretadas como indicadores, “síntomas” de deficiencias o excesos. Así, en lugar de interpretar las puntuaciones de los tests como “muestras de conducta”, se han interpretado como “signos” o “síntomas” representando un salto inferencial no justificado.
4. En cuanto a la utilización social, los tests comenzaron a usarse de forma generalizada para la clasificación de alumnos y/o selección de personal, lo que promovió la generación de una serie de tests que evalúan los aspectos potenciadores de la eficacia humana y no solamente las deficiencias (como se acentuaba en el psicodiagnóstico debido a sus raíces y sensibilidad médico-tradicional). En este sentido los tests psicológicos poseen, originariamente y debido a su mayor compromiso con una tradición diferencial (que no psicopatológica), una mayor carga en los elementos positivos de la psicología humana que la tarea propia del psicodiagnóstico.
5. El afán objetivista de la testología, obligada a convertirlo todo en números, produjo una corriente de psicólogos en la que todo debía expresarse en números para que fuera aceptado como científico, y otro grupo de profesionales que manifestaron un rechazo radical a todo lo que significara expresión numérica. En la polémica entre psicodiagnosticadores cualitativistas y testadores cuantificadores objetivistas, se abrió paso con fuerza la expresión de evaluación.

En toda tarea evaluadora se aplican tests o técnicas de evaluación cualquiera que sea el objeto de estudio. Por ello se produjo una equiparación de la Evaluación psicológica con los procedimientos de recogida de información a través de los cuales se realiza, y más concretamente con los tests, llevando incluso a denominar a esta disciplina "Tests psicológicos o Testing" (Fernández-Ballesteros, 1992a).

Testing hace referencia a un procedimiento mecánico, tipificado, cuantitativo, y objetivo, mientras que "evaluación" representa algo más, comprende desde el planteamiento de problemas hasta la emisión de juicios, un proceso mucho más complejo que el testing. El término medida o medición psicológica posee una significación menor, y por ello, entraría a formar parte, de una y otra manera, de la evaluación. Este término se utiliza en el contexto de la testología para referirse o bien a las operaciones de medida en los tests psicológicos fundamentalmente, o bien a los resultados de estas operaciones de medida que poseen una expresión numérica. Es este acento en lo cuantitativo, que en el contexto de la evaluación psicológica, lo diferencia del psicodiagnóstico en su concepción inicial.

Evidentemente el papel de las técnicas al aplicar el proceso evaluativo es grande. Ahora bien, el propio Cronbach (1990) señaló que la evaluación es un término más amplio que el correspondiente a la mera aplicación de tests, ya que connota la integración y valoración de la información recogida. Es por esto, por lo que en la actualidad ha decaído el término tests psicológicos o testing, incorporándose el de evaluación psicológica.

La evaluación psicológica se asentó en sus inicios en la medida, y ésta fue esencial para el avance de la propia psicología. Medir solo tiene sentido en un marco de relaciones. Sin tener un conocimiento amplio de ese marco, la medida carece de valor. Es ésta, quizás, la crítica que cabría realizar al "Testing" por el placer de medir y a las evaluaciones psicométricas carentes de objetivos específicos. Actualmente, todos estarían de acuerdo en aceptar que la evaluación psicológica no puede limitarse al uso de tests, debiendo incorporar otras técnicas como: entrevista, técnicas subjetivas, técnicas objetivas, técnicas psicofisiológicas, etc.

1.2.2. PSICODIAGNÓSTICO

Aunque el problema de los orígenes del Psicodiagnóstico es un tema complejo, parece que esta expresión fue utilizada por primera vez en 1921 por Rorschach, titulado de este modo el libro en el que presentó sus conclusiones sobre el test de las manchas de tinta.

El término diagnóstico etimológicamente deriva del griego *diá* - gignosko que significa aprender a conocer, conocer, tener conocimiento, entender algo... (Maganto, 1995b). Aunque en el origen este término no tuvo las connotaciones específicas que se han atribuido posteriormente (pedagógico, psicológico, médico...), en el campo de la psicología el concepto Psicodiagnóstico ha tenido unas connotaciones clínicas, en parte porque ha sido derivado de la medicina, y en parte porque ha estado vinculado a las perspectivas psicodinámicas.

La definición dada del Psicodiagnóstico en su acepción tradicional hace referencia al sentido médico-psiquiátrico del término ya que se plantea como objetivo exclusivamente la clasificación taxonómica. Es un diagnóstico dirigido a la exploración de aspectos psicopatológicos, negativos o deficitarios del funcionamiento psicológico, y la aplicación de instrumentos tiene por finalidad complementar, corroborar, y medir aquellas tendencias que, según la experiencia práctica de los clínicos, son de mayor significación.

En su análisis histórico y etimológico del concepto psicodiagnóstico, Pelechano (1988a) delimita 7 aspectos que dotan de significación a la expresión Psicodiagnóstico:

1. Se trata de una expresión con claras connotaciones médico-psiquiátricas.
2. Es un acercamiento empirista y aplicado en el que predomina la detección de diferencias entre los seres humanos identificados previamente por algún criterio externo.
3. Disminuye al máximo el momento discursivo teórico y se acentúan los correlatos operativos.
4. Pretende un estudio global del individuo acerca del funcionamiento y dinámica de su personalidad y de su inteligencia, lo que implica un acento en lo individual, por un lado, y por otro, un intento por plasmar la dinámica global, total del funcionamiento psicológico humano.
5. Acercamiento en el que prima lo cualitativo frente a lo cuantitativo.
6. Se trata de un método en el que se pretende conocer a un ser humano y se encuentra bien definida tanto la situación como los tipos de recursos que ha de tener el diagnosticador, el material, los modos de recoger las respuestas dadas por el psicodiagnosticado y los procedimientos concretos para llegar a obtener alguna clave interpretativa.
7. El método de trabajo implica una interacción verbal entre dos personas (psicodiagnosticador y psicodiagnosticado) y no-verbal, interacciones en las que se sugiere que no se influya en la persona que realiza la prueba.

Sintetizando algunos de los aspectos conceptuales del Psicodiagnóstico, Pelechano (1988a) indica que:

1. Se trata de un término acuñado por un médico (Rorschach) en el que la conducta observable no posee una entidad que se agote en sí misma como síntoma de algo.
2. Es un tipo de conocimiento que se encontraría situado entre la opinión y la ciencia, una suerte de arte en el que la experiencia personal y el modo de encadenar los conocimientos puede ser aprendido, aunque se asume la posibilidad de error.
3. Es un conocimiento en el que se acentúa lo cualitativo más que lo cuantitativo, lo global frente a lo analítico. Es el modo personal del diagnosticador y su experiencia, lo que aconseja el juicio definitivo.
4. Junto a una estimación acerca del estado actual, se trata de formular un juicio pronóstico acerca del curso que va a seguir la persona psicodiagnosticada en el futuro.
5. Es un tipo de actuación en donde predomina la observación y el registro, no el tratamiento ni la intervención directa. Se intenta identificar la alteración para posteriormente, llevar a cabo el tratamiento.

Este término estuvo caracterizado por un predominio de lo cualitativo sobre lo cuantitativo y se ha basado en gran parte en el conocimiento o la experiencia personal del diagnosticador que enfatizaba la importancia del individuo como tal, focalizándose sobre lo patológico. No obstante, aunque ciertamente la orientación psicoanalítica preferenció el término Psicodiagnóstico, así como Testing o medida en la corriente psicométrica, y evaluación en la perspectiva conductual, como subraya Avila-Espada (1986) no se encuentran razones suficientes para apoyar la especificidad o independencia de los distintos términos.

El término Psicodiagnóstico tuvo en sus inicios una clara semántica médica, pero tal denominación ha perdido en parte su vigencia, aunque en opinión de Fernández-Ballesteros (1992a, 2004a) este término sigue teniendo aplicabilidad cuando el objetivo de evaluación es el del diagnóstico psicológico: es decir, el de establecer la clasificación del sujeto en estudio con fines clínicos.

Entre las causas del declive del Psicodiagnóstico y la progresiva sustitución por el término Evaluación se pueden citar: 1) la apertura de nuevos ámbitos de trabajo; 2) la aparición de nuevas teorías que lo rechazan, como los enfoques fenomenológicos o los conductuales; 3) la asociación del psicodiagnóstico a las perspectivas psicodinámicas; (4) las críticas que plantearon la necesidad de buscar sistemas de evaluación acordes con los principios experimentales; y 5) las críticas sociales y éticas por el uso y abuso de los tests, potenció el rechazo de todo tipo de diagnóstico que se basara en ellos.

No obstante, cabe subrayar un resurgimiento del concepto, ya que en la última reunión del Área de Evaluación realizada en Madrid (22-9-2008), se decidió integrar de nuevo el concepto de “Diagnóstico Psicológico” junto con Evaluación Psicológica en la denominación de la asignatura para los nuevos planes de estudio que se desarrollarán dentro del Espacio Europeo de Educación Superior (EEES).

1.2.3. EVALUACIÓN

Se han vinculado los orígenes del concepto de evaluación a los trabajos llevados a cabo por Murray en Harvard Psychological Clinic. Estos trabajos tenían por objetivos: 1) descubrir las fuentes del funcionamiento personal creativo y adaptativo, y 2) estudiar a lo largo de un período largo de tiempo a un pequeño grupo de sujetos, explorando diversas áreas (desarrollo, historia escolar, desarrollo sexual, relaciones familiares...) mediante diversas técnicas (entrevista, autobiografía, tests objetivos y proyectivos, análisis de habilidades...), incluyendo observación natural y diseños experimentales especialmente concebidos para evaluar nivel de aspiraciones, resistencia al estrés, etc.

Durante la Segunda Guerra Mundial Murray y sus colaboradores prepararon un programa de evaluación para la Oficina de Servicios Estratégicos que tenía por objetivo predecir qué hombres podrían prestar un servicio eficaz. Por lo que se refiere a la expresión de evaluación (assessment) aparece por primera vez en la bibliografía como título de un libro *Assessment of Men* (O.S.S. 1948) en el que se recogen las actividades de la sección de psicología de la Oficina de Servicios Estratégicos de los Estados Unidos orientada por Murray. En este trabajo se opta por el término “evaluación” versus “Psicodiagnóstico”, debido a que este último insiste en la detección de elementos patológicos y/o defectuales, mientras que en “evaluación” se acentúan los aspectos positivos y de desarrollo potencial que existen en la conducta del ser humano. Aunque este término aparece por primera vez en este trabajo, el concepto se reafirma con la obra publicada por Mischel (1968) *Personality and Assessment*.

No obstante, el campo semántico que cubre “evaluación” se presenta desde sus orígenes como algo complejo y con diversas significaciones. Pelechano (1988a) en un intento de sistematizar-inventariar los principales tipos de significaciones que posee “evaluación” distingue tres: 1) Formación de impresiones y emisión de juicios acerca de los demás; 2) Un proceso, un modo de obtener información sobre una o varias personas, un modo de organizar esta información y una manera de interpretarla; y 3) Un proceso a lo largo del cual se toman

decisiones y se construye un modelo de trabajo sobre el funcionamiento y estructura psicológica de la persona que se trata de evaluar y de las situaciones que a esta persona le toca vivir.

1.2.3.1. EVALUACIÓN CONDUCTUAL

La apertura de nuevos campos (escolares, institucionales, terapéuticos, salud...), así como el surgimiento del enfoque conductual dieron pie a la sustitución del término Psicodiagnóstico por el de Evaluación conductual, modificándose y ampliándose el concepto. En la perspectiva de la evaluación conductual viene a tomar significación concreta la relación de diagnóstico y tratamiento, lo que debe ser un diagnóstico por y para la intervención, y así mismo un diagnóstico que no se agota en la antesala del tratamiento, sino que acompaña en todo momento... y lo sobrevive, pues existe un importante momento en cuanto al diagnóstico de postratamiento y de seguimiento.

Los términos "proceso y toma de decisiones" se recogen en la mayoría de las definiciones que los autores han propuesto (Blanco Picabia, 1986; Godoy y Silva, 1992; Kanfer y Saslow, 1965; Mischel, 1973; Pelechano, 1976, 1988a; Fernández-Ballesteros, 1992a, 1994; 2004, Maganto, 1995b) y en ellas se contempla siempre la recogida de información, la integración de la misma, así como la planificación de la intervención.

Con la publicación del famoso artículo de Kanfer y Saslow (Behavioural Analysis, 1965) se introduce en la evaluación la perspectiva operante del aprendizaje y la perspectiva experimental, que enriquecen una evaluación que hasta entonces había utilizado casi exclusivamente la perspectiva correlacional. Estos autores proponen un esquema que se basa en los 7 puntos siguientes:

1. Análisis de la situación problema, es decir, definición de los excesos o déficits conductuales, así como análisis topográfico de la conducta problema (intensidad, duración, frecuencia...).
2. Clarificación de la situación problemática que identifique las variables que mantienen la conducta problema.
3. Análisis motivacional, es decir, de los refuerzos y de los estímulos aversivos que pueden ser asociados con la conducta problema.
4. Análisis del desarrollo o de factores biológicos, sociales y conductuales que influyen en la conducta problema.
5. Análisis del autocontrol, es decir, situaciones o personas que refuerzan el autocontrol de la conducta problema.
6. Análisis de las relaciones sociales o de las personas que influyen positiva o negativamente en la conducta problema.
7. Análisis del ambiente sociocultural y físico.

Ya he mencionado lo que parece ser la principal contribución del análisis conductual: introducir la perspectiva experimental, lo que conlleva relacionar íntimamente las tareas diagnósticas con las de tratamiento. Con el enfoque conductual surge la preocupación por la eficacia de las intervenciones y la contrastabilidad de sus procedimientos. Surge así la evaluación conductual como un proceso de obtención de información que permite tomar

decisiones encaminadas a producir en los sujetos un cambio de conducta y la evaluación de este cambio, todo ello en el contexto de la terapia de conducta.

Los psicólogos conductistas en un principio estudiaron únicamente variables observables con el fin de establecer las relaciones funcionales entre las conductas y las variables que las mantienen. Pero posteriormente, las aportaciones de la corriente cognitivo-conductual promovieron que progresivamente se aceptaran variables mediacionales del mundo interno que no son directamente observables, siendo incorporadas éstas como objeto de estudio.

La Evaluación Conductual surgió en los años sesenta, en el contexto de la psicología del aprendizaje, como una alternativa al diagnóstico basado en entidades nosológicas y características individuales, y como respuesta ante las necesidades surgidas a partir de la modificación y terapia de conducta. Hasta los 80 se presenta como un claramente configurado marco de referencia, teórico y metodológico, pero tras estos 20 años se tronca en desilusión. A partir de los 90 aparece como un capítulo más dentro de la Evaluación Psicológica o bien como un ámbito de actuación.

La crisis de la Evaluación Conductual, en opinión de Fernández-Ballesteros (1994c) en la década de los 80, parece causada por el enfrentamiento de dos importantes fuerzas: la primera de ellas queriendo mantener planteamientos conductuales radicales, la otra tratando de compatibilizar el modelo conductual con otros enfoques cognitivos y personológicos. En su análisis, la autora destaca las asunciones conductuales que han sido olvidadas (análisis funcional y observación como único método) y las que han sido primero rechazadas y luego aceptadas (los autoinformes, variables personales explicativas). Sin embargo, en su opinión, la evaluación conductual sigue manteniendo un ámbito propio y específico, así como una presencia estable en las publicaciones psicológicas.

Desde esta perspectiva, Fernández-Ballesteros (1994c) define la Evaluación Conductual como aquella alternativa a la evaluación psicológica a través de la cual se trata de identificar las conductas objeto de estudio, tanto motoras como fisiológicas o cognitivas, así como las variables ambientales y/o personales que las mantienen o controlan, con el objetivo de realizar un tratamiento o cualquier tipo de intervención conductual o valoración.

1.2.3.2. EVALUACIÓN PSICOLÓGICA

El término evaluación psicológica surgió como consecuencia del enfoque conductista de la evaluación, connotando algo más amplio que la simple aplicación de tests, conllevando procesos de integración y valoración de la información recogida (Cronbach, 1990; Fernández-Ballesteros, 1992a, 2004a), y carente asimismo de las connotaciones analítico-clínicas del término psicodiagnóstico.

En las últimas décadas, el término Evaluación Psicológica ha irrumpido en el campo de la Psicología, considerándola a la vez una disciplina y un área de práctica profesional dirigida al análisis científico de la conducta humana y de los contextos en que ésta se desarrolla y cuyo objetivo es llegar a su descripción, nosología, comprensión, explicación, predicción y/o cambio. Así mismo se considera al sujeto de análisis activo interactuante y normalmente capaz de interregulación. El proceso evaluador se entiende como ayuda hacia el logro de un mejor desempeño en tal interregularización.

Siguiendo la línea de definición de nuestra materia iniciada por profesores representantes de diversas Universidades Españolas en Septiembre del 1981, Forns (1985) conceptualizó la evaluación del siguiente modo: "La evaluación psicológica es un proceso de análisis científico

de la conducta humana, este proceso de análisis conlleva un doble eje, diacrónico y sincrónico en relación al individuo objeto de análisis, atendiendo a la par: a) al sujeto como ser biológico y socio-cultural; b) a sus diversos ambientes condicionantes y reactivos ante su conducta; y c) a sus propias conductas en las tres variantes de motora, fisiológica y cognitiva. El objetivo de este proceso de análisis es llegar a una descripción, nosología, comprensión, explicación, predicción y/o cambio de la conducta del sujeto, que consideramos activo, interactuante y normalmente capaz de auto-regulación; tal proceso es entendido como ayuda (acción de tutela) hacia el logro de un mejor desempeño autoregulatorio”.

Algunos matices diferenciales entre Psicodiagnóstico y Evaluación Psicológica pueden clarificarse al analizar las conceptualizaciones realizadas por Fernández-Ballesteros. Esta autora en 1983 definió el Psicodiagnóstico como aquella disciplina de la psicología que se ocupa del estudio científico del comportamiento, de un sujeto o grupo de sujetos determinados, en su interacción recíproca con el ambiente físico y social, a fin de describir, clasificar, predecir, y, en su caso, explicar su comportamiento. No obstante, si bien el objetivo prioritario de la Evaluación Psicológica ha de ser el sujeto, conviene recordar que la evaluación puede dirigirse también a contextos, situaciones, grupos humanos, y a la valoración de intervenciones. Estos nuevos campos de aplicación de la evaluación psicológica presentan características específicas, por lo que se aplicará otro modelo a estas nuevas aplicaciones del Psicodiagnóstico (Fernández-Ballesteros, 1983a).

Posteriormente, Fernández-Ballesteros (1992a) definió la evaluación psicológica como aquella disciplina de la psicología científica que se ocupa de la exploración y análisis del comportamiento (a los niveles de complejidad que se estime oportunos) de un sujeto humano (o grupo especificado de sujetos) con distintos objetivos básicos o aplicados (descripción, diagnóstico, selección/predicción, explicación, cambio y/o valoración) a través de un proceso de toma de decisiones en el que se encaminan la aplicación de una serie de dispositivos, tests y técnicas de medida o/y evaluación.

Y recientemente, Fernández-Ballesteros (2004a) ha definido la evaluación psicológica como aquella disciplina de la psicología científica que se ocupa de la exploración y análisis del comportamiento (a los niveles de complejidad que se estime oportunos) de un sujeto humano (o grupo) con distintos objetivos (descripción, diagnóstico, selección/predicción, explicación, cambio y/o valoración de los tratamientos o intervenciones aplicadas). La evaluación implica un proceso de solución de problemas y toma de decisiones que comienza con una demanda de un cliente o sujeto (o grupo) a un psicólogo. Este procedimiento implica una serie de actividades científicas y profesionales, realizadas en una secuencia establecida con las que se pretende dar respuesta a esa demanda, entre otras, la administración de una serie de test, técnicas de medida y procedimientos de recogida de información.

Un análisis de los contenidos implícitos subyacentes a la utilización de ambos términos, permite observar que en ocasiones se asocia el término evaluación psicológica al proceso, mientras que psicodiagnóstico enfatiza el producto final de ese proceso, es decir, sus resultados (Echeburua, 1996c). No obstante, tendencialmente ambos términos se utilizan como sinónimos.

En más de una ocasión se ha descrito la evaluación como un acercamiento multirrasgo (puesto que se refiere a más de una dimensión psicológica humana), multimétodo (puesto que se utiliza más de un método de obtención de información), multifuente (recoge información procedente de más de una fuente) y/o multijueces (lo usual es que participe más de un evaluador a lo largo de todo el proceso). Se trata de un concepto típicamente multidimensional.

La evaluación precede al estadio final de valoración correspondiente a la toma de decisiones (por ejemplo, confirmar, modificar o terminar un programa de entrenamiento). Pero la evaluación no es totalmente independiente de la valoración puesto que aquella debe plantearse previendo posibles decisiones alternativas. En el proceso de evaluación se distinguen dos fases: 1) identificación del rango de conductas y atributos o características que se consideren relevantes y deben ser medidas, estimadas o evaluadas; y 2) selección y/o diseño de técnicas fiables para estimar, medir o evaluar tales conductas, rasgos, atributos o características.

Durante mucho tiempo se ha constatado una confusión entre lo que es la “aplicación de tests psicológicos” (*psychological testing*) y lo que es “evaluación psicológica” (*psychological assessment*). La evaluación psicológica es una tarea compleja que utiliza información derivada de tests, junto con los datos del historial, motivo de consulta, observaciones directas e informaciones de terceros (Rodríguez Sutil, 2007).

Como dicen Meyer et al. (2001), la aplicación de tests psicológicos es un proceso relativamente simple mediante el cual se aplica una escala concreta para obtener una puntuación específica (por ejemplo, la aplicación de un subtest de las escalas de Wechsler). Posteriormente se puede aplicar un significado descriptivo a esa puntuación, a partir de los hallazgos normativos, nomotéticos. La evaluación psicológica, en cambio, se ocupa del clínico que toma una serie de puntuaciones de tests, obtenidas generalmente de múltiples métodos, y considera los datos en el contexto de la historia, la información con que se derivó al paciente, y la conducta observada para comprender a la persona que está siendo evaluada, con objeto de responder a las preguntas que provocaron la derivación y comunicar los hallazgos al paciente, a sus personas cercanas y a la fuente de derivación. La evaluación psicológica tiene a su disposición una serie de instrumentos, cada uno de ellos con sus ventajas e inconvenientes que pueden ser aplicados en conjunto, de manera complementaria, en lugar de intentar determinar qué método de evaluación es, en general, el más exacto (Ganellen, 2007). El proceso de la evaluación psicológica se caracteriza entre otras cosas por ocupar un espacio prolongado de tiempo. No se ajusta desde luego a la cultura de los tiempos, propia del *fast food* y la solución inmediata (Rodríguez Sutil, 2007).

La diversidad de estrategias de evaluación pone de manifiesto las diferencias en la conceptualización de lo que entendemos por evaluación psicológica y por psicodiagnóstico. De algún modo el propio psicodiagnóstico puede ser un paso previo en el proceso de evaluación psicológica, entendiendo ésta como una actividad más exhaustiva y que implica más niveles de funcionamiento. La actualización de esta controversia o diversificación en la conceptualización de estos términos ha sido realizada por diversos autores (Affonso, 2005; Captao, 2005; de Fátima, 2007).

La evaluación psicológica es una actividad imprescindible que ayuda a la correcta identificación de los objetivos de intervención. Sin embargo, en salud mental es necesario que dicha evaluación se complete con otras técnicas y recursos que provienen de otros ámbitos científicos como la neuropsicología, por ejemplo, utilizando mapas cerebrales o técnicas de neuroimagen para cubrir los objetivos de una evaluación clínica completa en este ámbito (Vieira, da Silva, Neiva, 2007). En otros ámbitos de la evaluación psicológica, como por ejemplo en el ámbito de la evaluación de las discapacidades, es necesario la incorporación de modelos y técnicas de evaluación psicosocial para dar respuesta a los problemas surgidos en la propia evaluación. Una revisión actualizada de la evaluación psicosocial en el proceso de evaluación psicológica ha sido realizada por diversos autores (Fletcher y Greene, 2002; Levitas, 2001;

Poindexter, 2002; Sturmey, 2004, 2007). Cada vez más profesionales asumen un modelo de evaluación psicosocial para la evaluación y toma de decisiones respecto a la capacidad mental de las personas con problemas en esta área, desarrollándose instrumentos de evaluación innovadores que dan respuesta a esta necesidad (Brody, 2005; Newberry y Pachet, 2008; Pachet, Newberry y Erskine, 2007)

1.2.4. VALORACIÓN

En el libro de Adams (1964) no aparece el término de evaluación (*assessment*) sino el de valoración (*evaluation*) y medida (*measurement*). Para este autor la evaluación tendría una significación similar a medida y, en todo caso, es algo distinto de valoración: la valoración va más allá de la medición por cuanto supone juicios de valor. Partiendo de este presupuesto distingue 5 pasos en el proceso de valoración: 1) Determinación de lo que se quiere valorar; 2) Definición operativa de lo que se desea valorar; 3) Selección de las situaciones adecuadas que permitan obtener resultados; 4) Registro de las observaciones previstas de tal modo que sea susceptible de análisis objetivo; y 5) Resumen y síntesis de los datos recogidos y elaborados.

Sin embargo, no existe un claro consenso entre los autores para diferenciar evaluación de valoración. La valoración lleva consigo una toma de decisión y enjuiciamiento acerca del fenómeno o persona que se esté valorando, que no necesariamente requiere expresión cuantitativa, aunque sea deseable, y que entre estos enjuiciamientos, se considera deseable que se proponga más de un curso de acción para el futuro.

Evaluación y valoración son dos términos etimológicamente idénticos en castellano, sin embargo, el estudio empírico realizado por Fernández-Ballesteros (1992a) puso de relieve que evaluación (*assessment*) hace referencia, fundamentalmente, al examen o exploración de personas, mientras que valoración (*evaluation*) implica semejantes actividades referidas a objetos (puestos de trabajo, tratamientos, programas). Para Fernández-Ballesteros (1992g) por valoración se entiende la sistemática investigación del mérito, valor o éxito de concretos tratamientos, intervenciones o programas, llevados a cabo con el fin de producir cambios comportamentales en el contexto social. No obstante, esta autora posteriormente (Fernández-Ballesteros, 1995bc) utilizará el concepto evaluación de programas.

Así mismo, existen autores españoles (Anguera, 1994a; Fernández-Ballesteros, 1995b; Fernández del Valle, 1995; Forns, 1990, 1993; Forns y Gómez, 1995; Hernández, 1995; Martorell, 1988b; Sánchez y Anguera, 1993) que utilizan el término evaluación cuando hablan de programas, es decir, emplean ambos términos como sinónimos. Tal vez esta progresiva sinonimización de *evaluation* y *assessment* (valoración y evaluación) se produce para evitar utilizar la palabra valoración por su carga subjetiva relacionada con la emisión de juicios de valor.

Muchas de las aplicaciones en valoración de programas se han operativizado en el campo educativo, contexto en el que se han formulado una serie de condiciones para que estas valoraciones sean aceptables. Pelechano (1988a) indica algunas de las principales condiciones que deben tener estas valoraciones:

1. Deben proporcionar información suficiente como para poder tomar decisiones alternativas, lo que sugiere una pluralidad de enfoques y análisis.
2. Los resultados de la valoración deben permitir mejorar el programa que se está realizando o se haya realizado.

3. La valoración debe tener en cuenta los objetivos del programa inmediatos, a medio y a largo plazo.
4. Deben tenerse en cuenta tanto los objetivos del programa y los efectos estrechamente relacionados con ellos, como los efectos no esperados, ofreciéndose explicación tanto de los efectos esperados y los no esperados.
5. Los diseños de análisis a utilizar deben ser principalmente multivariados.
6. Los procedimientos de obtención de información deben cumplir los requisitos mínimos de objetividad, fiabilidad, validez, viabilidad, utilidad, y responsabilidad ética.

Si bien las tareas que realizan los psicólogos cuando evalúan y cuando valoran son prácticamente similares, el objeto y la finalidad perseguida varían. Los procesos de valoración de programas de intervención en salud, educación o servicios sociales son, como queda dicho, más complejos, y el psicólogo suele ser uno más en el equipo multiprofesional.

1.3. HACIA UNA PLURALIDAD CONCEPTUAL EN LA EVALUACIÓN PSICOLÓGICA

En la misma dirección que la señalada por Forns (1993) hace más de una década, en este proyecto se entiende la evaluación psicológica desde una perspectiva histórico-integradora, desde la que se considera que la evaluación psicológica incluye en parte los objetos de estudio que se han perfilado en los distintos momentos históricos y en distintas escuelas de la psicología, ampliándose el objeto de estudio de las variables intrapsíquicas, a las conductuales y de estas a las interaccionales, además de ampliarse su campo de actuación (clínica, escuela, organizaciones, judicial...), con las modificaciones conceptuales y metodológicas que ello ha implicado a la tarea de evaluación. Por ello, en este apartado se perfilan los cambios históricos operados en los objetivos de análisis de la tarea del evaluador que han conducido a una evaluación psicológica plural conceptual y metodológica en la actualidad.

En la primera década del S. XX el objetivo de la evaluación psicológica fue el análisis del sujeto en los aspectos biofisiológico y psicológico. Las características del sujeto se entendían como signos de sus *variables intrapsíquicas*, considerándolas consistentes a través del tiempo, estables a través de las situaciones, y determinantes de sus acciones presentes y futuras. En esta época la evaluación estaba íntimamente relacionada con los métodos de detección de las diferencias psicológicas que la psicometría desarrollaba progresivamente. Posteriormente, se realizaron análisis de los datos en términos de medias, porcentajes, distribuciones, estudios de los fenómenos según su ordenación ascendente, así como comparaciones ligadas a la edad. Después siguió el análisis de los fenómenos correlacionales y los análisis factoriales. Sobre las bases de esta disciplina psicométrica se asentaron los trabajos de los pioneros en la construcción de tests psicológicos. Aunque se dan importantes diferencias entre las escuelas, los modelos psicodinámicos, los enfoques humanistas fenomenológicos, y las concepciones del rasgo se centraron en el análisis cualitativo y cuantitativo de variables intrapsíquicas.

Aproximadamente, en la década de los 30 se va configurando el modelo conductista que en contraposición considera que el objeto de estudio de la evaluación psicológica deben ser las *manifestaciones externas del sujeto*, aquellas que son identificables mediante la observación. Estas conductas son consideradas como muestras del repertorio conductual del sujeto. En sus orígenes el análisis de la conducta se centró en el establecimiento de las relaciones funcionales entre los estímulos y la conducta del sujeto. Los evaluadores limitaban sus análisis a lo

observable y medible sin utilizar categorías de análisis que no pudieran operativizarse. Se potenció el uso de la metodología observacional y la creación de diseños de caso único. Inicialmente se estudiaron variables motoras, pero posteriormente fueron incluidas como objeto de análisis variables psicofisiológicas y cognitivas.

Este enfoque centrado en estudiar la conducta externa y sus relaciones con estímulos ambientales, dio pie al *estudio del contexto*, tomando progresivamente el ambiente más relevancia como objeto de estudio. El contexto fue definido en función de unas dimensiones bipolares: contexto presente-pasado, externo-interno al sujeto, molar-molecular, real-percibido... vinculándose a las conductas del individuo. En esta etapa los esfuerzos se centraron en la identificación de las variables que definen el contexto. Las observaciones de las características ambientales se factorializaron y se realizaron estudios para analizar las relaciones entre las variables ambientales y las conductas de los sujetos. Esta perspectiva de análisis estimuló la creación de escalas de evaluación de características físicas y estructurales del ambiente, así como de cuestionarios de percepción ambiental.

A partir de la década de los 60 comenzó a recuperarse el estudio del individuo como objeto de análisis, debido al surgimiento de la psicología cognitiva. Comenzaron a analizarse *los procesos cognitivos o procesos psicológicos internos*, subrayándose la importancia de estudiar los modos de percibir, pensar, memorizar... del sujeto. La psicología cognitiva produjo una gran influencia en los enfoques conductistas que incluyeron variables cognitivas y emocionales entre sus objetivos de análisis, después de aceptar la existencia de variables intermedias entre estímulo y respuesta. Desde la perspectiva de la psicología cognitiva el sujeto se considera como ser activo, capaz de manipulación y de representación de símbolos, en constante diálogo interaccional con la realidad que le circunda, y capacidad para tratar la información de su entorno y con mecanismos de feedback para autorregularse. Esto implica algunas transformaciones en el campo de la evaluación psicológica y las regulaciones y mecanismos por los cuales el sujeto interpreta la realidad se convierten en objeto de análisis. Desde esta perspectiva se aportan algunas nuevas técnicas tales como: técnicas de identificación de los componentes cognitivos, técnicas de análisis de los potenciales evocados, sistemas de organización de constructos cognitivos, técnicas de creación de disonancia cognitiva a través del contraste informativo en tareas de categorización de información, y métodos de análisis de tareas y de simulación de resolución de problemas. Además de estas nuevas técnicas continúan utilizando otras tales como la verbalización de pensamientos en voz alta, cuestionarios autoinforme o medidas sociométricas.

También desde la década de los 60 se desarrolla un enorme interés por el estudio y *evaluación de acciones psicosociales o acciones de planificación de tratamientos* para el grupo o la comunidad. Esto dio lugar al desarrollo de sistemas y técnicas de valoración de programas. La valoración de programas, como ya se ha señalado previamente, se basa en el supuesto conceptual de que la conducta de un grupo social puede ser modificada como resultado de unas acciones psicosociales debidamente planificadas que se ejercen de modo consciente y directo sobre los contextos en los que se pretenda crear un cambio.

La evaluación psicológica ha experimentado a lo largo de su historia importantes cambios conceptuales y metodológicos (Forns, 1993). Así, en su fase inicial, se confirió gran valor a las variables internas del sujeto como elementos determinantes de su actuación (posiciones intrapsíquicas). Se expandió después hacia el estudio de la presión del contexto próximo sobre la acción del sujeto, considerando a éste como pasivo-reactivo más que como agente de su

propia realidad; y las manifestaciones conductuales de los sujetos fueron entendidas como respuestas, observables y específicas, frente a estímulos determinados (posiciones behavioristas). Más tarde se reconoció la actividad del sujeto agente activo en la construcción de su propia realidad a través de un diálogo continuo con su entorno (posiciones interaccionistas). Finalmente se acentuó el papel de los contextos en los que se produce la conducta dando importancia al estudio del contexto y de los climas sociales. Recientemente, en un amplio orden socio-psicológico, el interés de la medida ha recaído en la valoración de resultados de las propias acciones psicosociales (evaluación de programas y de tratamientos), y en un orden psicológico, el análisis se ha centrado en el estudio de las operaciones cognitivas internas precisas para la operatividad humana (capacidad de planificación, autogüía, estrategias de resolución ante problemas....).

Para Forns (1993) la evolución que ha caracterizado la evaluación psicológica es de tipo centrífugo, copernicana (del interior del sujeto como punto central, al análisis de sus contextos) e integrativa. En momentos distintos, se ha acentuado un aspecto sobre otro y en el ámbito de la comunidad científica ninguna posición ha desplazado por completo a la anterior. Tal evolución ha sido pues expansiva a partir del individuo, y ha incorporado el análisis de la realidad en un triple sentido:

- Como externa al sujeto: análisis de las contingencias externas como entidades que por su relación funcional con el individuo creaban cambios o mantenían la conducta del sujeto.
- Como interactiva y constructiva con el individuo: proceso de mutua determinación en la dialéctica entre sujeto y objeto.
- Como sustantiva en sí misma: análisis de contextos y análisis de la incidencia de las acciones psicosociales - tratamientos y programas ejercidos en el exterior - sobre la conducta de los individuos.

En la misma dirección que Forns (1993) considero que la evaluación debería abordar simultáneamente en un modelo explicativo de la conducta, el análisis de las variables internas de un individuo; los determinantes fisiobiológicos personales; las conductas manifiestas y los sistemas de respuestas o reacción; los procesos cognitivos que subyacen en el actuar y el pensar humano; sus mecanismos de autocontrol y de planificación; las interacciones funcionales entre sujeto y realidad; el estudio sustantivo de los contextos y la eficacia de los contenidos y diseños de acción psicosociales planificados para ejercer un cambio en la conducta humana.

El objeto de la evaluación psicológica en la actualidad debe ser el estudio, análisis y valoración de las características del sujeto, de sus formas de acción, reacción e interacción con los demás y con la realidad, y de sus procesos de cambio. Todo ello analizado en el contexto de un sistema que está configurado por sujetos (bio-psico-sociológicos), sometidos a unos procesos (internos-externos) que afectan y determinan funcionalmente las formas de contacto del sujeto con la realidad (y viceversa) y cuyos efectos se expresan en productos o manifestaciones conductuales (motoras, fisiológicas, cognitivas y emocionales). Como ha señalado Forns (1993) todas las variables influyentes están sometidas a mecanismos de covariación o de interrelación recíproca.

La aceptación de la complejidad de la conducta humana implica atender a sus múltiples factores potencialmente causales. Esto conduce a un tipo de análisis que pone de relieve los distintos valores o pesos causales que reciben los elementos potencialmente explicativos

incluyendo acontecimientos situacionales, variables fisiológicas y genéticas, cambios en las variables, etc. Implica también atender a posibles efectos causales interactivos, a la forma que toman las relaciones funcionales y a las relaciones bidireccionales entre las alteraciones de conducta y los factores causales.

La psicología actual tiende a aceptar que la conducta es un sistema con un entramado de variables, cuya influencia es interaccional y dinámica, por lo que la medida de tal sistema debe conllevar el análisis del conjunto de sus elementos. Las formas de análisis deben ajustarse a la búsqueda de la explicación de la conducta en forma de estructura o de organización, siendo los métodos de análisis multivariados adecuados a este objetivo.

1.4. RESUMEN: EL CONCEPTO DE EVALUACIÓN PSICOLÓGICA

El término “*Evaluación Psicológica*” se ha ido imponiendo hasta prácticamente ser el dominante en el área de la literatura diagnóstica y/o evaluadora y/o testológica, formando parte de la denominación más frecuente en las monografías más recientes (Alonso Tapia, 2004; Fernández-Ballesteros, 2004; Garaigordobil, 1998; Moreno, 2003).

Una conceptualización reciente de ***Evaluación Psicológica*** (Fernández-Ballesteros, 2004a) la define como *aquella disciplina de la psicología científica que se ocupa de la exploración y análisis del comportamiento (a los niveles de complejidad que se estime oportunos) de un sujeto humano (o grupo) con distintos objetivos (descripción, diagnóstico, selección/predicción, explicación, cambio y/o valoración de los tratamientos o intervenciones aplicadas. La evaluación implica un proceso de solución de problemas y toma de decisiones que comienza con una demanda de un cliente o sujeto (o grupo) a un psicólogo. Este procedimiento implica una serie de actividades científicas y profesionales, realizadas en una secuencia establecida con las que se pretende dar respuesta a esa demanda; entre otras, la administración de una serie de test, técnicas de medida, y procedimientos de recogida de información.*

En la actualidad, bajo el rótulo de evaluación psicológica se ha incorporado una visión abierta e integradora, que considera a nuestra disciplina como un proceso de evaluación y que amplía el objeto de la evaluación psicológica a contextos, programas o intervenciones (Moreno, 2003). Forns, Abad, Amador, Kirchner y Roig (2002) subrayan que la evaluación de personas pretende:

1. Identificar la conducta central (problema o conflicto) del sujeto.
2. Relacionar esta conducta con variables que la desencadenan, mantienen o modulan, sean antecedentes o consecuentes, y expresar la relación por medio del diseño de un modelo hipotético de funcionamiento.
3. Diseñar un proceso de análisis psicológico, con finalidades descriptivas, comprensivas, predictivas, y/o explicativas del entramado de conductas que intervienen en el modelo, con el fin de validar la consistencia del modelo supuesto.
4. Seleccionar las técnicas de medida adecuadas al objetivo de análisis, y proceder a su aplicación.
5. Establecer el diagnóstico o toma de decisión diagnóstica: fijando el tipo de alteración referido a menudo a los sistemas taxonómicos DIM-IV o CIE-10; y describiendo el funcionamiento de las conductas alteradas en relación con todas las variables intervinientes.

6. Determinar el pronóstico.
7. Prevenir, predecir, orientar y aconsejar, ayudar a la formación, indicar el tratamiento.
8. Analizar la eficacia de las acciones de consejo, prevención, formación y/o tratamiento emprendidas.
9. Comunicar la información diagnóstica obtenida, y también la derivada del tratamiento, al paciente, persona o institución que ha solicitado la consulta psicológica.

Desde su perspectiva, *“habitualmente, el objeto de la evaluación es el análisis de la persona, de sus conflictos o problemas, pero la evaluación psicológica no se limita únicamente a este aspecto evaluativo. Las acciones humanas llevadas a cabo con el objetivo de hacer cambiar la conducta personal o social pueden ser también objeto de evaluación. Así la evaluación de tratamiento o la evaluación de programas forman parte de la acción evaluadora”* (Forns et al., 2002).

En líneas generales, se considera **“Psicodiagnóstico o Diagnóstico Psicológico”** cuando el objetivo de la evaluación sea establecer la clasificación del sujeto con fines clínicos, es decir, cuando el objetivo sea el diagnóstico psicopatológico de un individuo. Mientras que la **“Evaluación Psicológica”** es un constructo más amplio, que incluye al psicodiagnóstico y que se define como aquella disciplina que explora y analiza el comportamiento de un sujeto o un grupo de sujetos, con distintos objetivos (descripción del comportamiento, diagnóstico clínico, selección de personal, predicción de conductas, explicación de un comportamiento, cambio experimentado por una persona que ha realizado un tratamiento, y/o valoración de los efectos de una intervención) a través de un proceso de toma de decisiones en el que se emplean una serie de dispositivos (entrevistas, tests, técnicas observacionales, proyectivas, subjetivas, objetivas...), tanto para la evaluación de aspectos positivos como patológicos o clínicos. Tanto la evaluación del ambiente como la evaluación de programas son extensiones actuales de la Evaluación Psicológica y por lo tanto son incluidos bajo esta acepción.

Así, *la evaluación psicológica, entendida en un sentido amplio, abarcaría desde la descripción del comportamiento hasta la valoración de programas de intervención*. Tanto el psicodiagnóstico como la evaluación psicológica implican un proceso de recogida de información, y de organización y procesamiento de la misma sobre una persona con la finalidad de comprender, predecir y/o modificar su conducta, estando este proceso entroncado a la Psicología básica de la que toma sus teorías y conocimientos empíricos.

Para finalizar se puede destacar que en la última reunión de docentes de la asignatura, celebrada en la Universidad Autónoma de Madrid, el 22 de Septiembre de 2008, se enfatizó la necesidad de incluir dentro de la denominación de la asignatura el término diagnóstico, y se sugiere como nueva denominación de la materia **“Evaluación y Diagnóstico Psicológicos”**.

CAPÍTULO 2. HISTORIA DE LA EVALUACIÓN PSICOLÓGICA

- 2.1. Introducción
- 2.2. Fuentes del Psicodiagnóstico y de la Evaluación psicológica: Antecedentes precientíficos
 - 2.2.1. Míticos
 - 2.2.2. Racional-Especulativos
- 2.3. Proceso de constitución del Psicodiagnóstico (S.XIX) y desarrollos del S. XX en la Evaluación Psicológica
 - 2.3.1. Estructuración de Anastasi (Anastasi, 1968/1978; Anastasi y Urbina, 1998)
 - 2.3.2. Estructuración de Fernández-Ballesteros (Fernández-Ballesteros, 1992, 1994, 2004)
 - 2.3.3. Estructuración de Pelechano (Pelechano, 1981, 1988)
- 2.4. Notas sobre la crisis y resurgimiento de la Evaluación Psicológica
 - 2.4.1. Causas de la crisis del psicodiagnóstico tradicional
 - 2.4.2. Indicadores del resurgimiento de la evaluación psicológica
- 2.5. Evaluación psicológica: Últimos desarrollos
- 2.6. Resumen: Recorrido histórico de la evaluación psicológica

2.1. INTRODUCCIÓN

Una vez analizado el concepto es imprescindible, antes de pasar a ocuparnos de otras cuestiones, hacer un estudio de los orígenes de la Evaluación Psicológica. El análisis histórico de cualquier disciplina resulta una guía enriquecedora ya que proporciona una visión evolutiva de los avances tecnológicos acontecidos en la misma, así como de las concepciones asociadas a ella. Como han señalado algunas figuras relevantes en éste área "la perspectiva histórica añade un grado de profundidad al estudio de un tema particular, y mucho es lo que se puede aprender, tanto de los errores como de los aciertos de las investigaciones previas (Eysenck y Eysenck, 1985).

La historia de la evaluación psicológica, incluyendo el uso de los instrumentos de evaluación y las diferentes conceptualizaciones habidas ha sido recogida en varias de las revisiones realizadas por diferentes investigadores (Boyle, 2006, 2008; Boyle y Saklofske, 2004; Matthews y Anton, 2008). En Alemania una excelente revisión histórica ha sido realizada por Sprung y Sprung, 2001)

En este mismo sentido, investigadores preocupados por la evaluación reconocen que en épocas anteriores se han realizado importantes contribuciones. Por ello, en este capítulo se presentará una panorámica de lo que ha sido el desarrollo histórico de esta disciplina, lo que por otro lado nos permitirá ubicar los diferentes referentes teóricos y los instrumentos de evaluación. En este capítulo, en primer lugar se exponen las fuentes gestoras de la evaluación psicológica, posteriormente se analizan los principales desarrollos producidos como una disciplina de la psicología científica y, en última instancia, se clarifica su situación actual.

2.2. FUENTES DEL PSICODIAGNÓSTICO Y LA EVALUACIÓN PSICOLÓGICA: ANTECEDENTES PRECIENTÍFICOS

A pesar de que existen diversos trabajos que han estudiado los antecedentes del Psicodiagnóstico y/o de la Evaluación Psicológica, no existe acuerdo entre los autores respecto a las fases o períodos que permita una clara visión evolutivo-integradora. Esta situación se agrava más aún en el caso de que intentemos estudiar épocas anteriores al S. XIX. El motivo radica en el hecho de que en los períodos míticos y filosóficos, la psicología no tiene una ubicación concreta, ya que las aportaciones le llegan tanto de la Filosofía como de la Medicina o el Arte... Esta dispersión de fuentes, junto con la dificultad de su localización, hace que ese rastro sea sinuoso.

Se han realizado diversas sistematizaciones, y al analizarlas podemos constatar que cada investigador adopta distinto eje de estudio, algunos operativizan el desarrollo histórico a la luz de los modelos que sucesivamente han ido surgiendo en la Psicología, mientras que otros estructuran la secuenciación a la luz de los desarrollos técnicos, etc. Al ser varias las opciones, en este trabajo se opta por seguir un planteamiento cronológico genérico.

Al realizar un rastreo histórico, la primera pregunta que surge es ¿Se puede localizar el comienzo de la Evaluación Psicológica en un momento específico determinado de la historia del hombre? Pese a que han sido muchos investigadores los que se la han formulado, no es posible hallar una respuesta clara, debido a la dificultad de precisar el momento en el que por primera vez el ser humano comprueba la existencia de diferencias de cualquier tipo entre él y sus semejantes. Precisamente, es a partir de esta toma de conciencia cuando se buscan

procedimientos que permiten determinar la cuantía o el grado de esa discrepancia, y en este momento se podrían situar los orígenes de esta disciplina.

Si bien no ha sido posible conseguir un acuerdo en torno al primer antecedente de la evaluación, sí existe conformidad en cuanto a los requisitos que debería reunir para ser considerada como tal Evaluación. Estos requisitos serían: 1) Necesidad de un contexto y un objetivo para desarrollar dicha acción, es decir, necesidad de un evaluado, un evaluador, una técnica o procedimiento de medida y una interpretación de los valores obtenidos; 2) Existencia de un marco teórico que permita la ubicación de las hipótesis; y 3) Disposición de unos sistemas categoriales que permitan la clasificación del evaluado a partir de los resultados obtenidos con los procedimientos que se le hayan aplicado.

Aclarados estos aspectos realizaremos a continuación un análisis histórico-evolutivo del Psicodiagnóstico y/o la Evaluación Psicológica. Siguiendo la sistematización elaborada por Fernández-Ballesteros (1983b, 1992a, 1994c) podemos distinguir dos tipos de antecedentes precientíficos, por un lado, los antecedentes míticos, y por otro lado, los racional-especulativos.

2.2.1 MÍTICOS

La astrología ha sido considerada como el primer antecedente del Psicodiagnóstico por diversos autores (Fernández-Ballesteros, 1983b; Pelechano, 1988b; Silva, 1982). En concreto el horóscopo y las cartas astrales fueron las primeras herramientas con las que se trató de describir a las personas, así como de explicar y predecir su conducta. Como señaló McReynolds (1975), en la astrología se encuentran las características presentes en todo modelo evaluativo: 1) existe un sujeto a evaluar, un evaluador y unos protocolos, siendo el evaluador el que los interpreta; 2) existe un marco referencial teórico de la interpretación; y 3) existe un sistema de categorías taxonómicas en base a las cuales se realiza la explicación, predicción y descripción psicológica de los sujetos.

2.2.2. RACIONAL-ESPECULATIVOS

Dentro de los antecedentes se pueden destacar dos grandes fuentes de influencia en el Psicodiagnóstico, la Filosofía y la Medicina. Se pueden incluir aquí las aportaciones debidas a distintos filósofos y pensadores que con sus especulaciones han ido facilitando, en unos casos, o delimitando, en otros, toda la gama de cuestiones que posteriormente se ha planteado la Psicología y el Psicodiagnóstico. Entre las influencias que se han recibido de la Filosofía, tres aportaciones han sido especialmente relevantes, la de Aristóteles, la de Descartes, y la proveniente del Empirismo Inglés.

El planteamiento de Aristóteles es una fuente precursora del Psicodiagnóstico, por cuanto desde ella se propone el análisis del carácter y los hábitos psicológicos de las personas en base a sus características corporales. En esta perspectiva muchos filósofos se dedicaron a evaluar la forma de ser de las personas a través de sus rasgos corporales. La fisiognomía aristotélica deja sentir su influencia en los posteriores trabajos tipológicos de Kretschmer y Sheldon sobre el temperamento.

En segundo lugar, la influencia de Descartes desde cuya concepción se plantea que no es posible el estudio objetivo del hombre, por lo que hay que utilizar la intuición para acceder a los datos de la conciencia. Esta filosofía cartesiana subyace hoy en los enfoques

fenomenológicos y humanistas, desde los cuales se considera que no es posible una evaluación objetiva cuantificada. Pese a que esta concepción racionalista supone en parte un impedimento al desarrollo de esta disciplina, sin embargo, ha sido fuente de interesantes técnicas de evaluación (técnicas subjetivas).

Otra influencia es la del Empirismo Inglés desde el que se plantea la necesidad del análisis objetivo del hombre proponiendo como unidades de análisis las acciones, los hechos externos y objetivables. Esta filosofía subyacerá a las primeras formulaciones conductistas desde las que se consideró que el objeto de análisis debía ser la conducta externa del sujeto.

En esta fase, aunque no se hace ciencia tal y como hoy en día se entiende, sin embargo, se plantean las bases y prerequisites que van a facilitar el establecimiento de la posterior Psicología Científica.

Junto con la Filosofía, son las concepciones biológico-médicas las que dentro de la misma línea de pensamiento racional-especulativo influyeron sobre el Psicodiagnóstico. Cabe aquí destacar a Gall (1758-1828) como precursor de la evaluación neuropsicológica. Este médico planteó que las facultades psíquicas estaban localizadas en el cerebro y que en el análisis de las protuberancias craneales se podían hallar indicadores fieles de tales facultades. En base a estos presupuestos desarrolla la Frenología elaborando una serie de mapas del cráneo en los que localizó distintas características psicológicas como la emotividad o la inteligencia.

A partir de estas influencias se van perfilando tres perspectivas de la evaluación, una que enfatiza el estudio de los aspectos somáticos, otra que promueve el análisis de los fenómenos de la conciencia mediante la intuición y, en último lugar, otra corriente que planteará que son las manifestaciones externas, las actividades del sujeto, las principales unidades de análisis a la hora de realizar un estudio psicológico.

2.3. PROCESO DE CONSTITUCIÓN DEL PSICODIAGNÓSTICO (S. XIX) Y DESARROLLOS DE SIGLO XX EN LA EVALUACIÓN PSICOLÓGICA

Esta etapa, por ser la de mayor riqueza, productividad e identificación con los presupuestos científicos, se caracteriza por ser la más estudiada y sistematizada. Han sido muchos los autores que han tratado de listar los hechos y momentos esenciales de la historia del Psicodiagnóstico (Aiken, 1996; Carpintero, 1993; Fernández-Ballesteros, 1992a, 1994c; 2004a; Pelechano, 1988b).

Se han realizado diversas estructuraciones de este período de la historia del Psicodiagnóstico, con distintos enfoques. Entre los estudios aproximativos, algunos de ellos realizan un listado de hechos y momentos esenciales de la historia del Psicodiagnóstico, mientras que otros analizan la historia a la luz de los modelos y su productividad en relación a textos o herramientas desarrolladas por el enfoque. Esto da lugar a que la presentación de la historia que elabora cada autor sea en parte parcial. Por ello, y con la finalidad de obtener una visión más global e integradora de la historia, a continuación presentaré sintéticamente varias de estas estructuraciones debidas a relevantes autores como Anastasi y Urbina (1998), Fernández Ballesteros (1983, 1992a, 1994c), y Pelechano (1981, 1988) que nos permitan una visión comprensiva y totalizadora de esta disciplina.

2.3.1. ESTRUCTURACIÓN DE ANASTASI (Anastasi 1968/1987, Anastasi y Urbina, 1998)

La principal característica de la estructuración de Anastasi, en la que se presenta una seriación de momentos claves, radica en que ha sido realizada en base al auge y declive en el empleo de determinados instrumentos. Los períodos descritos por Anastasi no deben considerarse como sucesivos cronológicamente, sino como un intento de secuenciar diversas tendencias del Psicodiagnóstico. En la sistematización planteada por esta autora se proponen los siguientes períodos: 1) Período de las pruebas sensoriales y motoras; 2) Período de atención por lo mental; 3) Período de desarrollo de los tests de inteligencia; 4) Período de los tests colectivos; 5) Período de los tests de aptitudes; 6) Período de los tests tipificados de rendimiento; y 7) Período de la medida de la personalidad.

Período de las pruebas sensoriales y motoras: Puede situarse esta fase en el siglo XIX, siendo su rasgo principal la focalización sobre el estudio de los sentidos para llegar a un mejor conocimiento de los individuos y sus diferencias. Como principales representantes se identifica a Seguin, Esquirol y Galton. Además, se comienzan a analizar los resultados de estos primitivos tests con la ayuda de procedimientos estadísticos.

Período de Atención a la Mental: Los tests utilizados en este período, a pesar de su intención de evaluar "lo mental" son todavía eminentemente sensoriales (Binet y Henri, 1895). Así, el estudio de las diferencias individuales continúa centrándose en funciones como tiempo de reacción, sensibilidad al dolor... Simultáneamente desde la Psiquiatría comienzan a utilizarse con frecuencia procedimientos para el diagnóstico de los pacientes. Autores como J.M. Catell o Kraepelin pueden ser identificados como los más destacados de esta etapa.

Período del desarrollo de los tests de inteligencia: Este período se sitúa a comienzos del siglo XX siendo sus principales características: 1) Mayor interés por los procesos cognitivos superiores; 2) Diversificación de los estímulos incluidos en los instrumentos con el fin de facilitar el análisis de las distintas funciones; 3) Establecimiento de normas valorativas acerca de la edad cronológica de los examinados; y 4) Inicio del uso del concepto de Cociente Intelectual. En este contexto, las investigaciones de Binet y Simon, dieron pie a la creación del primer test de inteligencia. Anastasi destaca que el test de Binet-Simon con sus sucesivas revisiones atrajo el interés de los psicólogos de todo el mundo. Aparecieron traducciones y adaptaciones en muchos idiomas. En América se hicieron varias revisiones, de las cuales la más famosa fue la llevada a cabo bajo la dirección de L.M. Terman en la Universidad de Stanford, conocida como la revisión de Stanford-Binet (Terman, 1916).

Período de los tests colectivos: Como consecuencia de la entrada de los Estados Unidos en la I Guerra Mundial se incrementa el interés por la aplicación colectiva de instrumentos, así como un mayor empleo de tests objetivos y cuantificables a fin de calificar y clasificar a numerosos grupos de sujetos reclutas. La información obtenida con la aplicación de tests se utilizó en muchas decisiones administrativas, incluyendo la exención del servicio militar, la asignación a diferentes tipos de servicio o la admisión a los campamentos de formación de oficiales. En este período gran parte de los psicólogos norteamericanos se ponen al servicio del Gobierno y del Ejército, siendo su máximo exponente el Comité dirigido por Yerkes. En estas

circunstancias se constituyó el primer test de inteligencia colectivo. Con este fin los psicólogos del Ejército utilizaron diversos tests disponibles, así como un grupo de tests de inteligencia preparados por Otis. Los tests aceptados finalmente por los psicólogos del Ejército se han llegado a conocer como el Alfa y el Beta del Ejército. Sin embargo, para Anastasi, la aplicación de los tests de inteligencia colectivos no estuvo en relación directa con su perfeccionamiento técnico.

Período de los tests de aptitudes: La aplicación masiva de tests de inteligencia puso en evidencia que la expresión test de inteligencia era poco acertada puesto que solo medían ciertos aspectos de esta facultad, principalmente aptitudes verbales, y en menor medida relaciones numéricas, simbólicas y abstractas. Los profesionales de la Psicología toman conciencia de la necesidad de utilización de precisos instrumentos de aptitudes especiales que complementen los datos proporcionados por los denominados tests de inteligencia global. Mientras la aplicación práctica demostraba la necesidad de tests de aptitud diferencial, un desarrollo paralelo en el estudio de la organización de rasgos iba suministrando gradualmente los medios para su elaboración. Para ello se contó con la ayuda de procedimientos estadísticos y matemáticos (sobre todo Análisis Factorial), así como de teorías sobre la capacidad intelectual. Un ejemplo del papel que desempeñó el Análisis Factorial es la creación de Baterías de Aptitudes Múltiples. Estas baterías tienen por finalidad procurarnos una medida de la posición del individuo en cada uno de cierto número de rasgos. En lugar de una puntuación total o CI, se obtiene una puntuación por separado para rasgos, tales como la comprensión verbal, aptitud numérica, capacidad espacial, razonamiento aritmético y la velocidad de percepción. Las baterías de aptitud múltiple representan un desarrollo relativamente tardío en el campo de los tests. Casi todas han aparecido a partir de 1945. Autores como Spearman, o Thurstone pueden ser considerados los más característicos de este período.

Período de los tests tipificados de rendimiento: Mientras los psicólogos se ocupaban en desarrollar los tests de inteligencia, aptitud, los exámenes escolares tradicionales estaban experimentando ciertas mejoras técnicas. Los primeros tests tipificados para la medida de los resultados de la instrucción escolar comenzaron a aparecer a partir del comienzo de siglo. Encabezados por las investigaciones de E.L. Thorndike estos tests utilizaban los principios de medición desarrollados en los laboratorios psicológicos, Como ejemplos podríamos citar las escalas para estimar la calidad de la caligrafía y las redacciones, así como los tests de ortografía, cálculo y razonamiento aritmético. Más tarde aparecieron las baterías de rendimiento, con la publicación de la primera edición del Stanford Achievement Test en 1923. Prefigurando muchas de las características de los tests modernos, esta batería proporcionaba medidas comparables de la actuación en diferentes asignaturas escolares, valoradas en relación con un grupo normativo único. El origen de estos tests se puede ubicar en las demandas planteadas desde el terreno de la Educación. Durante esta fase se instituye en los Estados Unidos el empleo de baterías genéricas de tests, tanto en ámbitos estatales como federales, hecho facilitado por el inicio del empleo de correcciones mecanizadas. Sin embargo, los tests de rendimiento se utilizan no solamente con fines educativos, sino también para la selección de los aspirantes a puestos industriales y gubernamentales. A medida que un número creciente de psicólogos con conocimientos de psicometría participaba en la elaboración de los tests tipificados de rendimiento, los aspectos técnicos de los tests fueron pareciéndose cada

vez más a los de los tests de aptitudes e inteligencia. Los procedimientos para la elaboración y la valoración de todos ellos tienen mucho en común. El esfuerzo creciente por preparar tests de rendimiento que midieran la consecución de objetivos educativos amplios contribuyó también a que su contenido se pareciera más al de los tests de inteligencia. En la actualidad, la diferencia entre estos tipos de tests se cifra principalmente en el grado de especificidad del contenido y en la medida en que el test presuponga un curso expreso de instrucción previa.

Período de medida de la personalidad: Otro aspecto del examen psicológico se refiere a las cuestiones afectivas o no intelectuales de la conducta. Comúnmente, los tests destinados a este fin se denominan tests de Personalidad. Bajo este encabezamiento se incluirán, pues, rasgos intelectuales, sin embargo, en la terminología de los tests psicológicos, la designación de test de personalidad suele referirse a la medida de características tales como la adaptación emocional, las relaciones sociales, la motivación, los intereses, las actitudes... Anastasi considera como un precursor de los tests de personalidad a Kraepelin, quien utilizó el test de asociación libre con pacientes. En este test se da al sujeto palabras-estímulo seleccionadas y se requiere que responda a cada una exactamente con la primera palabra que se le ocurra. Esta fase se caracteriza por un mayor interés por aspectos cualitativos y un amplio enfoque del estudio de la personalidad. Se utilizan técnicas de asociación libre, cuestionarios de personalidad o self-report inventory, tests de ejecución o situacionales y técnicas proyectivas. Los autores más destacados abarcan una amplia gama tanto cronológica, profesional como metodológica e incluye a figuras como Kraepelin, Bleuler, Sommer, Galton, Pearson, y J.M. Catell, aunque también se pueden destacar a Woodworth, Hartshorne, May y Rorschach en calidad de representantes de los estadios iniciales.

La estructuración de Anastasi es adecuada y original, pero incompleta, ya que no incluye lo que son las nuevas etapas que están vinculadas con avances más recientes, es decir, los períodos de la evaluación conductual y el análisis funcional de la conducta, estudios psicofisiológicos y psiconeurológicos, el replanteamiento cognitivista y finalmente la psicología de la intervención.

2.3.2. ESTRUCTURACIÓN DE FERNÁNDEZ-BALLESTEROS (1983, 1992, 1994, 2004)

Desde otra óptica, Fernández-Ballesteros (1983b, 1992a, 1994c, 2004a) realiza un análisis amplio de los momentos histórico-evolutivos por los que ha ido discurriendo el Psicodiagnóstico. Esta autora parte, como ya se ha indicado en apartados anteriores, de un contacto de subordinación del Psicodiagnóstico con respecto a la Psicología. De este modo lleva a cabo una excelente revisión de las fuentes (míticas, racional-especulativas, científicas) que han ido incidiendo en el Psicodiagnóstico a consecuencia de los avances de la Ciencia. Posteriormente destaca algunas figuras relevantes para el proceso de constitución de esta disciplina, y finalmente enumera los hitos más significativos acaecidos de 1910 a la actualidad.

Antecedentes científicos

Fernández-Ballesteros (1983b) señala que el fundamental antecedente científico del Psicodiagnóstico es la constitución de la Psicología como disciplina experimental que se produce durante el último cuarto del siglo XIX. Los trabajos de **Brentano, Ebbinghaus, Dilthey,**

Hall, James, Wundt, entre otros, van a ser fieles exponentes de que la Psicología merece el estatuto de ciencia.

La fecha de constitución suele fijarse en 1879 cuando **Wundt** en Leipzig crea el primer laboratorio experimental de Psicología con la finalidad de estudiar cuestiones psicológicas con la metodología propia de la investigación fisiológica de la época: tiempos de reacción, sensación auditiva y visual, psicofísica y asociación. Wundt estableció tres supuestos para el estudio de la conciencia humana: 1) Todo elemento de la conciencia se deriva directa o indirectamente de la sensación; 2) Toda idea compleja es combinación asociativa o aperceptiva de elementos más simples; y 3) Los fenómenos mentales más complejos, como el pensamiento, no pueden ser estudiados en el laboratorio. Desde el punto de vista de la Evaluación Psicológica sus aportaciones más relevantes fueron: 1) Contextualizó la Psicología en el marco de las ciencias naturales, como ciencia empírica, y por lo tanto separada de la Filosofía; 2) Postuló el método experimental para el estudio del psiquismo; y 3) En aras de la objetividad, limitó las observaciones a los procesos más elementales de la conciencia (sensaciones, sentimientos, voliciones simples) y utilizó instrumentos para las mediciones (cronometría mental).

Desde sus comienzos la Psicología presenta una marcada disparidad en cuanto a su objeto, erigiéndose diversos modelos psicológicos (estructuralismo, psicoanálisis, conductismo, psicología de la gestalt...) con premisas conceptuales diferenciales.

Constitución del Psicodiagnóstico y/o de la Evaluación Psicológica

Como señala Fernández-Ballesteros (1992a, 2004a) en general hay consenso entre los historiadores de la psicología reconociendo que son tres los autores que comparten la paternidad del Psicodiagnóstico: Francis Galton, McKeen Cattell y Alfred Binet.

Galton (1822-1911) padre de la Psicología Diferencial y pionero del funcionalismo americano se dedicó a comprobar la herencia de las características humanas y en 1883 publica su obra más importante *"Inquiries into Human Faculty and its Development"* punto de partida de los tests mentales. En 1884 montó un "Laboratorio Antropométrico" donde cualquier persona podía conseguir medir su estatura, peso, capacidad respiratoria, fuerza muscular, rapidez de golpeteo, capacidad auditiva, agudeza visual, capacidad sensorial discriminativa y otra serie de mediciones sensoriales, perceptivas y motoras (DuBois 1970).

McKeen Cattell (1861-1934) en 1890 acuñó el término "test mental" y creó una serie de pruebas psicológicas para la medida de la presión dinamométrica, la velocidad de rotación del brazo, la habilidad para percibir estimulación en la palma de la mano, la memoria etc. El objetivo de tales pruebas lo centró en la determinación del rango, exactitud y naturaleza de las facultades psicológicas más elementales. Rechaza abiertamente la introspección y mantiene como objetivo prioritario la utilización de procedimientos objetivos en el estudio de las ejecuciones específicas de los sujetos sin formular sobre ellas inferencias sobre variables o atributos subyacentes. Las técnicas de evaluación de funciones sensoriales, perceptivas y motoras, y el énfasis en la utilización de medidas objetivas son las principales aportaciones de McKeen Cattell a la constitución del Psicodiagnóstico.

Alfred Binet (1857-1911) traspasa el interés por las funciones sensoriales, perceptivas y motoras planteando que el objeto de estudio deben ser las funciones psicológicas superiores: "no son las sensaciones, son las facultades psíquicas superiores lo que hace falta estudiar, son

ellas las que juegan el puesto más importante y la Psicología individual deberá prestar mucha más atención a ellas (Binet y Henri, 1895). Para este objetivo propone la utilización de tests mentales con los que estudiar la memoria, la naturaleza de las imágenes mentales, la imaginación, la atención, la comprensión, los sentimientos estéticos y morales y hasta la fuerza de voluntad.

Es el primer autor que emite un concepto claro de “*diagnóstico psicológico*”, como sinónimo de evaluación de la inteligencia, diferenciándolo del examen médico (exploraciones físicas y fisiológicas) y de la evaluación pedagógica (análisis de los aprendizajes). En 1905 crea el primer test de inteligencia destacando que un test debe tener tres requisitos: tareas simples, aplicación rápida y resultados contrastables por otros observadores. Este primer test de inteligencia que realizó en colaboración con Simón, introdujo por primera vez el concepto de “edad mental” que dio la posibilidad de expresar la medida de la inteligencia por medio del C.I.

El interés por la evaluación de las funciones superiores, la consideración del examen psicológico como complemento de la exploración médica y de la evaluación pedagógica, y la construcción de la primera medida de evaluación de la inteligencia parecen suficientes desarrollos como para considerar a Binet como uno de los padres del Psicodiagnóstico.

Además de estos tres autores, otros comparten el mérito de la constitución del Psicodiagnóstico. **Thorndike (1874-1949)** con sus trabajos base de la Psicología del aprendizaje y con la construcción de pruebas psicopedagógicas. Ebbinghaus con la creación de una serie de tests de completamiento con los que trató de medir las habilidades de los sujetos para la asociación de elementos verbales en una combinación significativa. Desde una perspectiva diferente, **Freud** proponiendo la evaluación del inconsciente mediante técnicas de asociación, dará lugar al posterior desarrollo de las técnicas proyectivas. Por último cabe destacar las aportaciones de **Pearson y Spearman** con el desarrollo de técnicas estadísticas para la construcción de tests.

La psicología científica establecida en el último cuarto del siglo XIX da lugar a la constitución del Psicodiagnóstico como disciplina. Entre 1890 y 1910 se puede situar esta constitución definiéndose en ese momento como una disciplina destinada al análisis de la individualidad a través de la evaluación de características sensoriales, perceptivas y motoras así como de otras llamadas superiores, a través de una serie de técnicas que adoptan el nombre de “tests mentales”.

Desarrollos desde 1910

Fernández-Ballesteros (1983b, 1992a, 1994c, 2004a) plantea los desarrollos del Psicodiagnóstico acaecidos desde 1910 en forma de síntesis cronológica destacando numerosos hitos que han permitido la evolución de esta disciplina. Debido a que sus referencias son muy amplias se recogen las que tienen más frecuencia de aparición en la literatura.

- 1915 Aparece el “test de laberintos” de Porteus a través del cual se pretende la evaluación de capacidades mentales libres de influencias culturales.
- 1916 Primera adaptación americana al test de Binet que adopta el nombre de “Escala de Stanford-Binet para la medición de la inteligencia”.
- 1918 Se emprende la evaluación de la inteligencia del ejército americano a través de los tests “Army Alpha” y “Army Beta”.

- 1920 Woodworth construye el primer test colectivo de medida del ajuste personal, el "Personal Data Sheet" para la selección de los soldados americanos.
- 1919 Gessell edita su escala de evaluación del desarrollo de las primeras edades.
- 1921 Rorschach publica su "*Psychodiagnostik*" en el que presenta su técnica de evaluación de la personalidad mediante la interpretación de manchas de tinta.
- 1926 Goodenough edita el Test de Dibujo de la Figura Humana para la medida de la inteligencia infantil.
- 1929 Se celebra en Barcelona el Congreso Internacional de Psicología en el que se dedican varias ponencias a temas relacionados con la evaluación psicológica.
- 1934 Moreno publica su obra "*Who Shall Survive?*" en la que, con el planteamiento de las técnicas sociométricas, se inicia la evaluación de interacciones sociales, y se presenta el "role-playing" como técnica de medida.
- 1935 Vygotsky formula el concepto de "zona de desarrollo próximo", antecedente de la importancia de la evaluación de estrategias cognitivas frente a productos intelectuales.
- 1936 Son abolidos los tests mentales en la Unión Soviética como instrumentos de medida.
- 1936 Se publica la primera escala de medida del desarrollo social, la "*Vineland Social Maturity Scale*" de Doll.
- 1938 Murray presenta su teoría sobre la personalidad (en colaboración con Morgan) junto con un instrumento proyectivo de evaluación, el Test de Apercepción Temática.
- 1939. Wechsler publica la primera edición de su medida de la inteligencia.
- 1942 Hathaway y McKinley presentan el primer inventario clínico de evaluación de entidades nosológicas psiquiátricas, el MMPI "*Minnesota Multiphasic Personality Inventory*".
- 1947 Halstead publica "*Brain and Intelligence*" lo cual supone un importante puntal en la evaluación psiconeurológica.
- 1949 Tyler, en su "*Basic Principle in Curriculum and Instruction*", sienta las bases para la valoración de programas e intervenciones.
- 1950 Shapiro presenta a la Reunión Anual de la Sociedad Británica de Psicología el trabajo "*An experimental approach to diagnostic psychological testing*" en el que se sientan las bases metodológicas para una evaluación experimental.
- 1952 Se constituye la Sociedad Española de Psicología.
- 1953 Skinner edita su obra "Ciencia y conducta humana" de la cual se deriva el análisis experimental de la conducta humana como alternativa conductista al Psicodiagnóstico en base a tests.
- 1957 Cronbach y Gleser publican "*Psychological Tests and Personnel Decisions*" en el que se propone utilizar la teoría de la decisión al proceso psicodiagnóstico.
- 1963 Staats se refiere, por primera vez, a la evaluación conductual en su texto "*Human Complex Behavior*".
- 1965 Kanfer y Saslow, con la publicación de su trabajo "*Behavioral Analysis*" plantean un nuevo modelo de Psicodiagnóstico desde una perspectiva conductual.
- 1968 Publicación del trabajo de Mischel "Personalidad y Evaluación" el cual representa la más dura crítica al modelo de evaluación en base a rasgos realizado a través de los tests.

- 1972 Feuerstein edita su *"The Dynamic Assessment of Retarded Performers"*, que supone el inicio de un modelo cognitivo en la evaluación del funcionamiento intelectual.
- 1974 Moos publica *"The social climate scales: an overview"* que representa el más importante esfuerzo en la evaluación del ambiente.
- 1976 Los soviéticos Voitko y Guilbuj presentan un importante trabajo "Acerca de ciertos conceptos principales en psicodiagnóstico" el cual significa una síntesis conceptual del modelo de la Psicología soviética.
- 1977 Se consolida la evaluación conductual con la publicación de numerosos textos tanto en el extranjero como en España.
- 1980-1990. Etapa de desencanto y crisis de la evaluación conductual.

El análisis de estos eventos históricos le lleva a Fernández-Ballesteros (1992a, 2004a) a establecer una serie de conclusiones. En primer lugar, que gran parte de los sucesos relatados son relevantes no sólo para la Evaluación sino para otras áreas o especialidades de la Psicología. Así, la Psicología Diferencial y la Psicología de la Personalidad ejercen un papel importantísimo, ya que van a significar nuestro sustrato metodológico y teórico. Así también, los avances de la Psicopatología, la Psiconeurología, la Psicología social, así como la aplicación de los principios de la Psicología del aprendizaje a la clínica, han supuesto hechos históricos determinantes en la construcción de esta disciplina. Por último, las exigencias sociales en el ámbito clínico, escolar y de las organizaciones han promovido la elaboración de técnicas evaluativas, dando lugar a la tecnología de la Evaluación psicológica.

En segundo lugar, de lo enumerado se deriva que la mayor parte de las realizaciones producidas en el ámbito de la Evaluación, hasta por lo menos los años 50 se han centrado en el desarrollo de una tecnología evaluativa dirigida al análisis y constatación de atributos psicológicos, entidades nosológicas y constructos dinámicos. Sin embargo, existen escasas elaboraciones sobre la formulación de las bases conceptuales y metodológicas de la propia disciplina, es decir, se echa en falta una reflexión sobre el objeto y el proceso de la Evaluación Psicológica. Es a partir de los 60 cuando comienza una nueva época en la que, junto con la puesta en cuestión y crítica de una evaluación exclusivamente realizada en base a tests de medición de características individuales, van desarrollándose modelos (como el conductual y el cognitivo) y tecnología conectada con la Psicología básica. Por otra parte, es también a partir de los años 70 cuando se extiende el propio objeto de la evaluación, pasando del énfasis en el análisis de las características individuales a la evaluación de interacciones persona-ambiente y, de ahí, a la evaluación del ambiente y a la valoración de intervenciones.

En tercer lugar, el análisis de los hechos históricos permite perfilar una serie de modelos: El modelo del atributo, el modelo dinámico, el modelo conductual, el modelo cognitivo y el modelo médico. Así, los desarrollos de la Evaluación, como disciplina derivada de la Psicología científica, se plantean cinco modelos fundamentales, los cuales, partiendo de distintos objetivos de evaluación y ámbitos de aplicación, van a dar lugar a concretos procedimientos y formas de hacer evaluación.

Analizando la estructuración planteada por Fernández-Ballesteros se pone de relieve la realización de una revisión muy completa de los hitos más significativos del desarrollo del Psicodiagnóstico y/o de la Evaluación Psicológica hasta finales de los 70. Sin embargo, apenas se perfilan los desarrollos de la evaluación psicológica en la década de los 80 y los 90.

2.3.3. ESTRUCTURACIÓN DE PELECHANO (1988)

Este autor hace un novedoso planteamiento del estudio de esa historia en el que se contemplan conjuntos de contenidos doctrinales homogéneos en cada período, pero simultáneamente relacionados con el contexto cultural de cada época y con las aplicaciones que de ellos se derivan. Pelechano (1988b) conceptualizando la psicología como una ciencia social, considera que intentar ofrecer un panorama histórico-racional de la Evaluación Psicológica resulta imposible sin una consideración de las implicaciones que lleva consigo conceptualizar a la psicología como una ciencia fundamentalmente social, así como sin hacer referencias a partes de la historia de aquellas disciplinas a partir de las cuales la Evaluación Psicológica se ha desarrollado, es decir, la psicología diferencial, la psicopatología, la psicología clínica, la psicometría, la psicología evolutiva, la psicología educativa, la psicología del trabajo... Así, desde esta perspectiva considera 5 momentos con características propias y específicas: 1) De la segunda mitad del S. XIX a 1921; 2) de 1921 a 1945; de 1946 a 1965; 3) de 1966 a 1975, y 4) una última etapa que abarca el final de los años 70 y los 80.

De la segunda mitad del S. XIX a 1921

1. *Sensibilidad romántica*: La sensibilidad romántica se caracterizaría, entre otras cosas, por un afán extremado en encontrar explicaciones a los hechos y acciones individuales y colectivas. Además, este sentir y pensar romántico, acentúa la consideración de la unicidad y/o individualidad del ser humano.
2. *Diagnóstico neurológico y psiquiátrico*: Por otro lado, hay que recoger los progresos que, desde la praxis clínica médica en general, neurológica y psiquiátrica en particular, fueron haciendo en las décadas anteriores, en relación naturalmente con la praxis diagnóstica.
3. *Evolucionismo biológico y darwinismo social*: En tercer lugar, considera el impacto de la teoría evolucionista dentro del pensamiento que dio lugar a una psicología diferencial y como pre-requisito, establecimiento de criterios para la evaluación de las diferencias individuales.
4. *Constitucionalismos*: El acercamiento constitucionalista de corte biológico recogía, en esta época, tanto las distintas corrientes fisiognomistas como el movimiento frenológico.
5. *La Psicofísica*: El desarrollo de la fisiología experimental y el problema de la relación alma-cuerpo se creyó resolver mediante un estudio sistemático y cualificado de estas relaciones.
6. *El Legado del Asociacionismo filosófico del empirismo inglés*: Desde Locke (1632-1704) el empirismo inglés se había caracterizado, entre otras cosas, por una tendencia a elaborar una filosofía práctica, preocupada por la búsqueda de la felicidad y la emisión de normas de vida encaminadas a su logro.
7. *Interés creciente en educación*: Se produce un creciente interés por la educación y la psicología educativa, tanto en Estados Unidos (Stanley Hall) como en Europa (Binet).

En síntesis, este período se caracteriza por desarrollarse en un ambiente cargado de sensibilidad romántica, en cuyo seno emergen los progresos de la Psicología en relación al incremento del interés por lo personal e individual, todo ello bañado por el impacto de las teorías evolucionistas y sus derivaciones en relación al interés por las diferencias cualitativas y por el incipiente interaccionismo. Simultáneamente, a este período corresponden los intentos de la caracterología y la aparición de medios que permiten el inicio de los estudios experimentales y objetivos de las diferencias individuales en el laboratorio. Como figura significativa del compendio de todas estas influencias se ubica aquí al polifacético Galton (1822-1911) y los cuatro aspectos de su obra, que se caracteriza por: Su interés por lo cuantitativo, la obtención de resultados específicos, el análisis de indicadores de las características que detectaba y su preocupación por los aspectos hereditarios de esas características (Pelechano, 1981). Entre las fechas y obras representativas de este período Pelechano (1988) destaca:

- 1869 Galton publica *Hereditary Genius*.
- 1884 Galton ofrece pruebas psicológicas en el South Kensington Museum.
- 1890 McKeen Cattell utiliza la expresión “test mental”.
- 1905 Binet publica la primera métrica de la inteligencia.
- 1916 Terman publica la primera adaptación de la escala de Binet en EEUU.
- 1921 Yerkes publica “Psychological Examining” en “The United States Army”.
- 1921 Rorschach publica “Psychodiagnostik”.

Desde 1921 a 1945

El conjunto de eventos históricos que se producen en este período resulta indicativo de la riqueza de acercamientos, la existencia de distintas escuelas de pensamiento a nivel teórico y técnico-metodológico, la efervescencia que promovió la creación de instrumentos y teorías que se encuentran vivas en la actualidad y, a la vez, la aparición de fuertes polémicas entre defensores y detractores de los distintos acercamientos. En este período Pelechano (1988b) identifica 5 características relevantes:

1. *Formulación del modelo psicométrico básico.*
2. *Potenciación del uso de tests y pruebas psicológicas.*
3. *Generación de una alternativa cualitativa de análisis del caso individual con pruebas proyectivas.*
4. *Dos polémicas centrales: estructura factorial de la inteligencia y psicometrista-proyectivismo.*
5. *Gran producción bibliográfica de tests y técnicas proyectivas.*

Dentro del acercamiento psicométrico (tests) es el período en el que surge una doble polémica, la que se refiere a las alternativas monofactoriales y multifactoriales en el estudio psicométrico de la inteligencia, y el enfrentamiento entre acercamientos cuantitativos (tests) y cualitativos (pruebas y técnicas proyectivas) en la evaluación.

Respecto a la polémica en torno a la estructura factorial de la inteligencia Spearman plantea la existencia de un factor general de la inteligencia, mientras que Thurstone defendía la existencia de una pluralidad de factores intelectuales, igualmente importantes, independientes entre sí, y responsables de la inteligencia humana. En la actualidad se acepta la existencia de un factor general de inteligencia que no llega a explicar toda la varianza y factores de grupo relativamente independientes, que cubren un porcentaje de varianza significativa. En relación a la segunda polémica se observa que como reacción al cuantitativismo psicométrico y despersonalizado de los tests, las pruebas proyectivas adquieren prestigio en la evaluación de la personalidad dentro de la praxis clínica.

En más de una ocasión se ha dicho que todo lo nocivo que tienen las guerras para la humanidad, lo tienen de bueno para el desarrollo científico y, al menos, por lo que se refiere al siglo XX, parece que esa afirmación no se encuentra muy lejos de la realidad. Física, medicina, química, sociología, economía, psicología..., han alcanzado desarrollos significativos a consecuencia del ritmo de trabajo impuesto durante los períodos de conflicto bélico internacional.

Un indicador de gran importancia en la evaluación psicológica al final de este período lo representa el papel desempeñado por los tests en la Segunda Guerra Mundial. En Estados Unidos, una comisión de psicólogos participaron en la evaluación y selección del personal militar. Este hecho estimuló que tanto en Estados Unidos como en Europa comenzaran a construirse pruebas psicométricas de la inteligencia y de la personalidad. Además, el fin de la guerra provocó una serie de problemas de integración de excombatientes a la vida civil. En Estados Unidos, cerca de 1500 psicólogos estuvieron prestando servicio como clínicos durante la contienda, tanto en labores de evaluación como de intervención terapéutica. A la vuelta de la guerra se va a plantear la valoración del diagnóstico psiquiátrico de corte tradicional y la posibilidad de generar un nuevo modelo de evaluación y de tratamiento terapéutico. El primer requerimiento promueve la polémica entre diagnóstico clínico y actuarial. La segunda opción genera el análisis de conducta o análisis funcional, como modelo alternativo de evaluación psicológica. Pelechano (1988b) enfatiza las siguientes obras representativas de este período:

- 1927 Strong publica el Inventario de Intereses Vocacionales.
- 1933 Berlin. Hitler nombrado canciller alemán. Se prohíben oficialmente los tests.
- 1935 Murray publica el Test de Apercepción Temática (TAT).
- 1936 Moscú. El Comité Central del Partido Comunista prohíbe el uso de tests.
- 1936 Doll publica la Escala de Madurez Social de Vineland.
- 1938 Bender publica el Test Gestaltico-visomotor.
- 1938 Thurstone publica "Primary Mental Abilities".
- 1942 Hathaway y McKinley publican el Inventario Multifásico de Personalidad de Minnesota (MMPI).
- 1945 Termina la Segunda Guerra Mundial.

Desde 1946 a 1965: Los comienzos de una crisis

1. *Predominio de los Estados Unidos como consecuencia de la emigración de científicos europeos debida a la Guerra:* Una guerra siempre representa una crisis fuerte para la sociedad que la sufre. Pero, también, representa una crisis social el final de la guerra; la tarea bien de reconstrucción nacional, bien de readaptación personal, económica, social, laboral, etc. se presenta como una acción de primera necesidad. La Segunda Guerra Mundial no fue una excepción. Representó una ruptura con la estructura y dinámica social, desaparición de instituciones de investigación, y por lo que se refiere a Europa (uno de los lugares físicos en los que se desarrolló el conflicto), la emigración masiva de científicos a Estados Unidos o Latinoamérica, práctica desaparición de ciencia psicológica de primera línea durante algunos lustros y, complementariamente, comienzo de un predominio progresivo de Estados Unidos en cuanto a ideas, instrumentación y modos de intervención ante los cuales los europeos, comenzarán por aprender primero, imitar después y, tan sólo a finales de esta época, en todo caso, comenzarán, poco a poco, a coger el pulso del progreso científico en temas de evaluación psicológica.
2. *Polémica en torno al uso social de los tests y a la entidad científica de las pruebas proyectivas:* Por un lado, la gran difusión de los tests y técnicas de evaluación psicológica, así como la práctica obligatoriedad de su utilización provoca un planteamiento crítico sobre los modos de actuación práctica, sobre los supuestos teóricos de los que parte y de la adecuación ética y moral en la utilización de este tipo de instrumentación para la clasificación y selección de personas en los mundos educativos, en el laboral, y hasta en el clínico. Por otro lado, se cuestiona la validez y fiabilidad de las técnicas proyectivas.
3. *Aparición de la Psicología Clínica:* Toma carta de naturaleza la psicología clínica como profesión y rama de la psicología en 1946.
4. *Aparición de la Terapia de Conducta con el correspondiente modelo de evaluación:* La aparición de este enfoque terapéutico representó un cambio importante en las concepciones clásicas de la evaluación. La Evaluación Conductual propone como objeto de evaluación únicamente lo observable.
5. *Cambios en criterios de valoración del sistema educativo y del sistema industrial:* Dentro de la psicología del trabajo, se comienzan a desarrollar técnicas de grupo encaminadas a modificar la conducta del sujeto que vive inmerso en un sistema organizacional; aparece el estudio sistemático de las relaciones humanas en la empresa y es la época, por lo demás, de fomento de la productividad en donde el ejecutivo representa un elemento clave dentro del sistema, es el hombre-organización. Además, es la era en la que se publican varios manuales clásicos dentro de la psicología del trabajo. En el ámbito educativo, los informes sobre valoración de programas comienzan a aparecer al final de este período. La preocupación por la educación lleva consigo la realización de múltiples tipos de trabajos acerca de la bondad del sistema en sus diversas partes: métodos de enseñanza, sistemas de valoración de los alumnos, métodos de registro y evaluación.

Entre los numerosos autores de este período se puede destacar a Eysenck y Catell por sus trabajos sobre la evaluación factorial de la personalidad. Kanfer, Saslow y Skinner con sus aportaciones al análisis funcional de la conducta. Así mismo, las aportaciones de Kelly, Osgood y Stephenson en relación a las técnicas subjetivas tienen un papel importante en esta etapa. Entre las fechas y obras más relevantes de esta época se pueden señalar las siguientes:

- 1946 Despegue de la psicología clínica.
- 1950 Gulliksen. Teoría de los tests mentales.
- 1951 Eysenck. Estudios de validación sobre eficacia terapéutica del psicoanálisis.
- 1953 Aparición de la Terapia de Conducta.
- 1954 Sociedad Americana de Psicología.
- 1959 Campbell y Fiske plantean la necesidad de una matriz multirrasgo-multimétodo para los estudios de validación.
- 1965 Kanfer y Saslow publican un modelo complejo de análisis funcional en la Revista Internacional de Psiquiatría.

Desde 1966 a 1975: La década imperial conductista

1. *Expansión del análisis funcional o análisis de la conducta:* En este período el conductismo se hace con la Psicología científica “oficial”, con autores que extienden el análisis funcional y la modificación de conducta a problemas muy distintos.
2. *En el modelo clásico refundición y búsqueda de nuevas opciones:* En este momento se reelaboran aspectos parciales y se detectan una serie de deficiencias que deberían ser superadas. Cronbach et al. (1972) formulan un modelo teórico omnicompreensivo acerca de la fiabilidad y validez de los tests psicológicos, el modelo de generalizabilidad, modelo teórico que no resulta, al final, integrado dentro de la praxis testológica debido, entre otras cosas, a su complejidad.
3. *Carta de naturaleza del situacionismo y pragmatismo:* Un momento decisivo del pragmatismo funcionalista americano plasmado en el conductismo es la apelación al situacionismo. Cobra relevancia el estudio de variables ambientales versus las variables de personalidad evaluadas con la testología. Este situacionismo critica tanto a los tests psicológicos como a las pruebas proyectivas. Este rechazo a la evaluación de variables de personalidad plantea: a) que son las situaciones las que determinan las conductas, y b) que para modificar las conductas el objeto de estudio debe ser el ambiente.
4. *Comienzan disensiones en la modificación de conducta y separaciones significativas:* Una nota central de la década que estamos comentando es la salida del conductismo del laboratorio y en esta salida se enfrenta a problemas sociales graves y relevantes, a nivel no solamente individual sino organizacional e institucional. Después del primer momento asociado al éxito aparecen los problemas asociados a la consolidación de las respuestas, efectos de la generalización... lo que provoca que hacia 1976 aparezca la polémica entre conductistas y cognitivistas.

5. *Reconversión de modificadores de conducta individual en intervencionistas comunitarios:* El énfasis en el ambiente y en lo social estimula la incorporación de modificadores de conducta skinnerianos al movimiento, recién creado, de la psicología comunitaria.
6. *Se pierde purismo de planteamientos teóricos:* Todas las posibilidades que han surgido suponen un planteamiento pluriforme y confuso en el que los autores parecen mucho menos adscritos a escuelas de pensamiento que a especialidades psicológicas y a elaboraciones de estas especialidades.

Entre los principales autores y obras de este período se mencionan:

- 1966 McReynolds. Avances en evaluación psicológica.
- 1968 Mischel. Personalidad y evaluación.
- 1968 Barker. Psicología ecológica.
- 1969 Bandura. Principios de Modificación de Conducta.
- 1970. Kanfer y Phillips. Fundamentos de aprendizaje de la terapia de conducta.
- 1972. Cronbach, Gleser, Nanda y Rajaratnam. Dependencia de medidas comportamentales.
- 1973. Wiggins. Personalidad y predicción.
- 1974. Davidson, Clark y Hamerlynch. Valoración de programas comportamentales.

El final de los 70 y la búsqueda de orden y sentido

El final de los años 70 representa una fase de espera-transición que posee las siguientes características:

1. *Momento de Crisis:* Dentro de la psicología contemporánea de estos últimos diez años parece que se vislumbra un cierto sentimiento de crisis y cansancio para todas las escuelas, paradigmas y modos operativos. Frente a los "istmos" y posiciones, en general, radicales de los momentos anteriores, al final de los setenta parece que la psicología ha alcanzado el nivel de tolerancia necesario como para poder valorar objetivamente los resultados que se han ido alcanzando.
2. *Valoración de programas:* Es el comienzo de la publicación de monografías, en las que aparecen en el título cada vez más frecuentemente la expresión de valoración tanto de la psicoterapia como de los programas de política social y educativa.
3. *Menor rigidez metodológica y doctrinal:* Al final ya de los años setenta se intenta hacer buena la expresión de que "lo mejor es enemigo de lo bueno", en el sentido de disminuir las exigencias metodológicas y procedimentales con tal de seguir teniendo contacto con problemas complejos que se dan en el mundo social y personal, pero que no pueden estudiarse de modo riguroso en el laboratorio.
4. *Incorporación de técnicas tradicionales:* Se observa la reincorporación de las técnicas y modelos de análisis estadísticos en el análisis funcional o análisis de la conducta. No deja de

resultar nuevo que un número del "*Journal of Applied Behavior Analysis*" dedicado explícitamente a exponer las señas de identidad del análisis de conducta contiene una terminología muy similar a la que se puede encontrar en los manuales "clásicos" de validación de tests dentro de la tradición psicométrica. Ello indica que al margen de la interpretación concreta que se haga de los resultados obtenidos, las técnicas operativas de la evaluación psicológica tradicional se encuentran incorporadas.

5. *Creación de nuevos instrumentos con diferentes objetivos de análisis:* Se siguen utilizando los tests psicológicos de inteligencia y los cuestionarios e inventarios de personalidad. En los últimos años de la década de los setenta y comienzos de los ochenta, dada la ampliación de nuevos campos en los que la psicología está colaborando, se van creando nuevos instrumentos, específicos para estos campos. Por otra parte, el impacto del procesamiento de información sobre partes considerables de la psicología (atención, memoria y pensamiento...) ha llevado tanto a replantear los instrumentos convencionales utilizados para la evaluación de estas parcelas como a realizar un reanálisis de los tipos de ítems y/o problemas que componían las pruebas de evaluación de estas parcelas, y reinterpretando los resultados en función de un compromiso diferenciado de distintos subprocesos-componentes, en un esfuerzo renovado por unir las pruebas psicológicas con esta alternativa teórica.
6. *Desarrollo de la Ecopsicología y la Psicología ambiental:* El grupo de Moos, desde 1974 ha publicado monografías acerca de la dimensionalización de distintos ambientes vitales-vivenciados en una línea con antecedentes tan notables como K. Lewin. Barker, por su parte, desde su monografía de 1968 intenta una dimensionalización física de los ambientes, contextos, escenarios o entornos dentro de los cuales tiene lugar la conducta de las personas y que determina un volumen de varianza igual o mayor que el correspondiente a la información entresacada de pruebas psicológicas convencionales.
7. *Críticas a la adecuación de los criterios de eficacia en la evaluación:* La proliferación de áreas estudiadas así como la participación de psicólogos en una serie de decisiones de política social (en calidad de miembros de comisiones de estudio) o miembros de comisiones encargadas de valorar los efectos que tales decisiones producen, ha llevado al planteamiento crítico de la adecuación de los criterios de eficacia aducidos de modo acrítico por los políticos. Por poner dos ejemplos ilustrativos: la aplicación de tests de inteligencia en la tercera edad y las calificaciones escolares como criterio de rendimiento educativo.

En definitiva, la última década ha acrecentado la complejidad de perspectivas. En ella ha aparecido un cierto relativismo respecto a la interpretación y significación de los resultados que se pueden obtener con distintos procedimientos de evaluación, se ha recuperado gran parte del modelo psicométrico clásico a la vez que se están incorporando nuevos procedimientos de análisis, ha desaparecido la posición xenófoba entre testólogos y analistas de conducta y las posiciones contextuales de ecopsicología y psicología ambiental van ganando posiciones.

Así mismo, aparece una psicología comprometida con los problemas sociales y a la vez que defiende tareas de profilaxis comportamental y de innovación social. Ello implica la reorientación de la investigación con relación a las épocas anteriores: una tarea importante de

la evaluación psicológica consiste en la creación de instrumentos y técnicas encaminadas a la detección de personas y/o comunidades con “alto riesgo”, que distan mucho de aquellas técnicas o instrumentos clasificatorios de antaño.

2.4. NOTAS SOBRE LA CRISIS Y RESURGIMIENTO DE LA EVALUACIÓN PSICOLÓGICA

Diversos autores han analizado las causas de la crisis del Psicodiagnóstico tradicional en la década de los 60-70, así como los hechos que indican un resurgimiento del mismo en la década de los 80. Con la finalidad de clarificar los factores destacados como más significativos, a continuación se presentan aquellas causas que con mayor frecuencia han sido enfatizadas en la literatura.

2.4.1. CAUSAS DE LA CRISIS DEL PSICODIAGNÓSTICO TRADICIONAL

Se suelen considerar las publicaciones de Mischel (1968) entre otras como el punto de partida de la crisis. Sin embargo, un análisis más detenido permite advertir que los problemas venían gestándose de antes, aunque es en la década 1965-1975 en la que se produce el punto álgido en la crisis. Al analizar las causas que determinaron la decadencia del denominado Psicodiagnóstico Tradicional se subrayan principalmente:

- Críticas a las Técnicas proyectivas: Se produce un desencanto por los datos negativos sobre la validez y fiabilidad de las técnicas proyectivas que eran las de mayor uso en el contexto clínico.
- Auge progresivo de la Evaluación conductual: Este enfoque inicialmente surgió tratando de ser una alternativa al Psicodiagnóstico Tradicional. Desde la Evaluación conductual se propuso la descripción de los problemas en términos conductuales versus en términos psicodinámicos, se plantean que la selección y evaluación del tratamiento debe ser el objetivo de la evaluación versus el etiquetado psiquiátrico y se enfatiza la utilización de la observación, la entrevista y los autoinformes versus los tests.
- Ataques a los tests: Otra de las causas que contribuyeron a la decadencia del Psicodiagnóstico fue el uso excesivo de los tests, que llegó casi a un abuso acompañado de una gran repercusión social. Este descrédito se centró fundamentalmente en la valoración de la inteligencia en el ámbito escolar y en interpretaciones basadas en tests proyectivos.

En relación a esta fase de revuelta antitests Blanco-Picabia (1993) considera 4 razones de la misma: 1) Las limitaciones reales tanto técnicas de los tests como humanas; 2) Los abusos por su utilización masiva e indiscriminada, así como por la utilización en ocasiones con fines ideológicos, políticos o comerciales, con las consiguientes situaciones éticas que plantean; 3) Las reacciones defensivas de la sociedad consciente de la existencia de personas dotadas de instrumentos y conocimientos que temen que pueda violar su intimidad o su libertad; y 4) Además, se plantean críticas derivadas de una inadecuación en la presentación de los resultados del proceso psicodiagnóstico.

Tal vez no sea casualidad el hecho de que durante el congreso de la Asociación de Psicología Aplicada en Montreal (1974) se creara una comisión, International Test Commission (ITC), con la finalidad de controlar la distribución y el uso de tests. Una revisión de la prehistoria de la Comisión Internacional de Test fue realizada por Cardinet (1995). Resulta difícil valorar la intensidad de la crisis. A pesar de todo lo que se ha insistido en ella, la crisis ha sido menor de lo que se ha predicado. En todo caso la decadencia que atravesaba el Psicodiagnóstico se detuvo y en la década de los 80 surgió un nuevo interés por los tests y la evaluación.

2.4.2. INDICADORES DEL RESURGIMIENTO DE LA EVALUACIÓN PSICOLÓGICA

Entre los indicadores más significativos del resurgimiento de la evaluación en la década de los 80 se pueden destacar (Calonge, 1987):

En primer lugar, numerosas encuestas realizadas en los EEUU confirman la *pervivencia de los tests tanto en su utilización práctica en la clínica, como en la enseñanza, en los programas de postgraduados...* Entre los tests más utilizados destacan el MMPI, el WAIS-WISC y el Rorschach. Estos resultados parecen relativamente concluyentes, a pesar de algunos defectos y sesgos evidentes. Sin embargo, todas las encuestas adolecen del defecto común a este tipo de indagación, a saber, que los resultados están viciados por el tipo de pregunta efectuada; es decir, ninguna de estas encuestas permite conocer en qué casos, situaciones, en qué momento de la evaluación y ni siquiera con qué objetivos, se utilizan unas y otras técnicas.

En segundo lugar, podemos tomar como indicador el *incremento de tiempo que se ha venido dedicando al Psicodiagnóstico y/o Evaluación Psicológica en las convenciones de la División 12 de la American Psychological Association (APA)*.

En tercer lugar se puede tomar como índice las *numerosas publicaciones* que aparecen referidas al Psicodiagnóstico o a sus instrumentos de evaluación más característicos: los tests. Se publican varios manuales de Psicodiagnóstico y Evaluación Conductual, apareciendo cinco volúmenes de *Advances in Personality Assessment*, editados por Butcher y Spielberger, que se suma a los *Advances in Psychological Assessment* editados por McReynolds previamente. Además, en 1989 comienza a editarse la revista denominada *Psychological Assessment. A Journal of Consulting and Clinical Psychology*, publicada trimestralmente por la *American Psychological Association*. Esta revista surgió tras una división de la *Journal of Consulting and Clinical Psychology* que se restringió a "artículos empíricos relacionados con el diagnóstico, tratamiento y prevención de los desórdenes clínicos, mientras que la nueva publicación se centra en artículos empíricos dedicados a la evaluación y valoración clínica.

En cuarto lugar, es sintomático de la vitalidad de la evaluación psicológica la *aparición de nuevos tests así como las frecuentes revisiones de tests tradicionales* (WAIS, MMPI). Esta actividad en el campo de los tests ha incluido también a los tests proyectivos. Un ejemplo de ello se ha operado con el Rorschach a través de las aportaciones de Exner en la línea de proporcionar tanto un sistema único de clasificación e interpretación como en la búsqueda de validación empírica de hipótesis proyectivas.

Además, en esta época se producen una serie de hechos que en parte pueden explicar esa revitalización de la evaluación psicológica:

1. *Disminución del Antagonismo entre la Evaluación Conductual y el Psicodiagnóstico Tradicional:* Cuando surge la Evaluación Conductual lo hace en medio de una crítica devastadora del diagnóstico tradicional, al que incluso parece que viene a sustituir. Sin embargo, a mediados de los 80 se pone de relieve que la evaluación conductual supone simplemente un enfoque de la evaluación psicológica, a la que ha hecho valiosas contribuciones. A esta aproximación entre ambos enfoques ha contribuido el enfoque conductual cognitivo que ha ofrecido recientes aportaciones a la evaluación.
2. *Aplicación de las Teorías de la Información a la evaluación de la inteligencia:* Nuevas teorías prenden y se desarrollan en el campo del Psicodiagnóstico creando una creciente base empírica. No deja de ser curioso que esta revitalización incida precisamente en el campo de evaluación de la inteligencia que habiendo sido pionero y central en la historia del Psicodiagnóstico, sin embargo, sufrió particularmente la devaluación que supuso la crisis.
3. *Desarrollo de importantes instrumentos en el campo de la Evaluación Psiconeurológica:* Aparecen instrumentos fundamentales como la revisión de la batería Halstead-Reitan o la de la batería Luria-Nebraska.
4. *La ampliación del objeto de estudio de la Evaluación Psicológica:* En el panorama actual de la Evaluación psicológica se amplía el objeto de estudio más allá del individuo, incluyendo parejas, grupos, organizaciones, contextos y ambientes, así como la valoración de intervenciones.
5. *La creación de nuevos campos:* Se crean nuevos campos y especialidades que abren nuevas perspectivas al Psicodiagnóstico, como son el campo de la Psicología de la Salud (Buela-Casal, Caballo y Sierra, 1996; Godoy, 1996; Roa, 1995), el judicial, el deportivo...
6. *La introducción de los ordenadores:* Se produce una eclosión vertiginosa que abarca la aplicación de todo tipo de tests así como su interpretación, la aplicación de entrevistas clínicas por ordenador y la elaboración de informes.
7. *Desarrollo de manuales de evaluación psicológica desde distintos enfoques o modelos:* Fisher (1985) propone una evaluación fenomenológica humanística, centrada más en el vínculo entre entrevistador y entrevistado con el objetivo de explorar el proceso de vida de la persona, y menos centrada en los tests, su administración y devolución. El texto de Wetzler y Katz (1989) planteado a partir de un enfoque behaviorista-cognitivo propone una evaluación más focal, basada en principios de observación conductual cuyo objetivo es la medición formal y rigurosa de los componentes psicopatológicos. Otros manuales plantean posturas integradoras de la evaluación tradicional con los nuevos aportes. Un ejemplo de este tipo es el libro de Swiercinsky (1985) que incorpora el DSM-III-R y al mismo tiempo integra categorías de análisis de raigambre más psicodinámica como fortaleza del yo, estilo de las defensas..., proponiendo así mismo combinar técnicas tradicionales con otras escalas más focales.

2.5. EVALUACIÓN PSICOLÓGICA: ÚLTIMOS DESARROLLOS

En la década de los 90 Watkins (1991) se preguntó por el estado de salud de la evaluación llegando a la siguiente conclusión: "Aunque el papel de la evaluación ha pasado por tiempos duros en el pasado, parece ser bastante resistente y estar experimentando un cierto renacimiento en la última década". Basándose en un metaanálisis de todas las encuestas publicadas entre los años 1960 y 1990 sobre la enseñanza y la práctica de la evaluación psicológica en el campo de la clínica en los Estados Unidos encuentra que: 1) existe una estabilidad en la práctica a lo largo de las tres décadas; 2) es una parte esencial de la formación del psicólogo en los programas de postgrado; y 3) un 50-75% de los psicólogos realizan evaluaciones y un 10% de ellos se dedica a esta tarea con exclusividad. Análisis del estado de la evaluación clínica en USA (Frank de Verthelyi, 1992) plantean que no sólo se ha mantenido vigente como afirma Watkins, sino que ha adquirido mayor respetabilidad y status científico tal como lo pronosticaban Korchin y Sundberg.

Reflexionando sobre la situación y las perspectivas futuras de la Evaluación Psicológica podemos considerar que esta disciplina tiene un largo pasado y un futuro lleno de retos. Fernández-Ballesteros (1995) en el III Congreso Nacional de Psicología Clínica y de la Salud concluyó que el proceso de evaluación es más que una estrategia racional elaborada teóricamente por los evaluadores, siendo un producto elaborado de la actividad epistémica del psicólogo cuando este tiene que tomar decisiones en relación con el sujeto de evaluación e intervención. Con vistas al futuro, la autora consideró que se iban a producir varios desafíos en distintas disciplinas, entre los que mencionó:

- La psicología de la salud en cuyo contexto se estaban desarrollando medidas relativas a tipos de personalidad (A y C) como predictores de patologías tales como las enfermedades cardiovasculares y el cáncer.
- La evaluación de la calidad de vida como concepto psicológico.
- Desarrollos importantes en el campo de la evaluación de programas.
- Estudio de las distorsiones de respuesta, y sobre las traducciones e importación sin adaptación de los tests psicológicos.
- Otra línea de futuro trabajo se relacionaba con ámbitos de evaluación desarrollados a partir de nuevas tecnologías como lo es la realidad virtual y su aplicación en multitud de áreas de la evaluación desde la capacidad espacial a los tests situacionales.

En España, la década de los 90 (Moreno, 2003), supone la creación oficial de las Facultades de Psicología, y el comienzo del nuevo siglo supone para la Evaluación una serie de retos al verse implicada tanto directa como indirectamente en el "Título de Especialista en Psicología Clínica" o en la integración en el marco del "Espacio Europeo de Educación Superior" (EEES). El empleo de los tests, tan frecuente unos años antes, pasa por un período de crisis, reflejo del conflicto entre Evaluación Conductual (EC) y Tradicional (ET), sin embargo, posteriormente la dicotomía entre la EC y la ET se supera, y en la actualidad se utilizan los tests clásicos trabajándose con ellos nuevas orientaciones. Además, se utilizan las Escalas de Evaluación, que emergieron con la Evaluación Conductual, llevando a cabo con ellas adaptaciones de las pruebas originales y no únicamente traducciones. Tal y como predijo Pelechano (1988) se ha producido una consolidación de escuelas y modos de distintos de

pensar la evaluación, en donde técnicas proyectivas, tests, y análisis de conducta se complementan con valoración de programas de intervención de índole distinta.

La situación de la evaluación desde la década de los 90 y hasta la actualidad se puede caracterizar por lo siguiente:

- La evaluación ha vuelto a ser incorporada como una actividad esencial dentro de los servicios de salud, muchos de los cuales testean rutinariamente a todos sus pacientes. Las preguntas se han hecho más puntuales y requieren a su vez respuestas más precisas. El psicólogo se ha convertido en un experto y un activo asesor, colaborando en la planificación del tratamiento y en su evaluación posterior.
- Una exigencia actual respecto al psicólogo es que realice esta tarea con el mínimo de tiempo y coste profesional. El tema del ahorro del tiempo-costo, está promoviendo un incremento en la utilización de software que restringe la interacción entre el psicólogo y el paciente.
- Otro objetivo de la evaluación clínica actual consiste en llegar a un diagnóstico ajustado al DSM-IV-TR.
- Empleo de múltiples técnicas (autoinformes, proyectivas...), así como actualización de tests tradicionales (MMPI-2, WAIS III...) y creación de nuevas técnicas.
- Los cuadros psicopatológicos más actuales son: depresión, trastornos de la alimentación, delincuencia, adicciones y los efectos disociativos de situaciones traumáticas.
- Reflexiones sobre el papel del género en los diversos tests y estudios transculturales en relación a los tests.
- Desarrollo de manuales de evaluación psicológica, entre los más recientes publicados cabe destacar: Tests psicológicos y evaluación (Aiken, 1996); Tests psicológicos (Anastasi y Urbina, 1998); Manual de evaluación en psicología clínica y de la salud (Buela-Casal, Caballo y Sierra, 1996); Pruebas y evaluación psicológica. Introducción a las pruebas y a la medición (Cohen y Swerdlik, 2001); Evaluación psicológica: historia, principios y aplicaciones (Gregory, 2001); Coordinadas, procesos y garantías (Alonso, 2004); Evaluación psicológica. Conceptos, métodos y estudio de casos (Fernandez-Ballesteros, 2004); Manual para la evaluación clínica de los trastornos psicológicos: estrategias de evaluación, problemas infantiles y trastornos de ansiedad (Caballo, 2005); Manual para la evaluación clínica de los trastornos psicológicos. Trastornos de la edad adulta e informes psicológicos, (Caballo, 2006); Pruebas Psicológicas: Principios, aplicaciones y temas (Kaplan y Saccuzzo, 2006).
- Incremento de la utilización del ordenador en el proceso de evaluación psicológica.

La evaluación psicológica ha evolucionado en las siguientes direcciones: 1) Continuado desarrollo de técnicas focales; 2) Creciente sofisticación psicométrica; 3) Constante actualización del DSM; 4) Mayor desarrollo de técnicas administradas por ordenador; 5) Mayor sofisticación de la interpretación por ordenador; 6) Mayor integración entre teoría y práctica; y 6) Más énfasis en la formación multicultural. No obstante, y a pesar de estas evoluciones, cabe resaltar la necesidad de incorporar el acelerado desarrollo científico y técnico, y responder a las crecientes demandas del medio, pero manteniendo como eje la importancia del vínculo humano.

Kaplan y Saccuzzo (2006) observan que durante las últimas décadas han surgido y prosperado varias de las principales ramas de la psicología aplicada: neuropsicología, psicología de la salud, psicología forense y psicología infantil, las cuales han ampliado el uso de las pruebas de psicológicas. De este modo se ha incrementado su uso y ha mejorado su estatus. Como en el pasado, la evaluación psicológica en la primera década del Siglo XXI continúa siendo uno de los tópicos más importantes, aunque controvertidos, en psicología. La evaluación es uno de los elementos fundamentales de la psicología. Aunque no todos los psicólogos utilizan pruebas y algunos se oponen a ellas, todas las áreas de la psicología dependen del conocimiento obtenido en los estudios de investigación que se basan en estos métodos de medición. En la compleja sociedad actual, la relevancia de los principios, aplicación y discusión de la evaluación psicológica se extiende más allá del campo de la psicología. Aún cuando un individuo no tenga planes de convertirse en psicólogo, es probable que se enfrente a pruebas psicológicas. Abogados, médicos, trabajadores sociales, gerentes de negocios, educadores y muchos otros profesionales deben tratar con frecuencia con informes basados en tales pruebas. Probablemente las pruebas nunca han sido tan importantes como lo son en la actualidad.

2.6. RESUMEN: EVOLUCIÓN HISTÓRICA DE LA EVALUACIÓN PSICOLÓGICA

En este apartado se sintetizan los aspectos más relevantes relativos a la historia de la evaluación psicológica (Cuadro 2) descritos a lo largo del capítulo.

Cuadro 2. Evolución histórica de la Evaluación Psicológica

ANTECEDENTES PRECIENTÍFICOS	
Míticos	La astrología
Racional Especulativos	Filosofía: Aristóteles (Fisiognomía), Descartes (Evaluación subjetiva) y el Empirismo Inglés (Evaluación objetiva) Medicina: Gall (Evaluación psiconeurológica)
PROCESO DE CONSTITUCIÓN DEL PSICODIAGNÓSTICO Y DESARROLLOS DE LA EVALUACIÓN PSICOLÓGICA DURANTE EL S.XX.	
Estructuración de Anastasi (Anastasi, 1968/1978/1988; Anastasi y Urbina, 1998)	
Período de las pruebas sensoriales y motoras Período de atención a lo mental Período de desarrollo de los tests de inteligencia Período de los tests colectivos Período de los tests de aptitudes Período de los tests tipificados de rendimiento Período de la medida de la personalidad	
Estructuración de Fernández-Ballesteros (1980, 1983, 1992, 1994, 2004)	
Antecedentes científicos	• Psicofísica: Weber y Fechner. • Avances de la Fisiología: MacDax y Broca. Burdach. Ferrier. • La psiquiatría organicista: Pinel. Kraepelin. Esquirol. • La pedagogía experimental: Stanley Hall • Otros: Brentano, Ebbinghaus, Dilthey, Hall, James, Wundt.
Constitución del Psicodiagnóstico	• Galton: Psicología Diferencial. • Mckeen Catell: Test mental • Binet: Primer test de inteligencia • Otros: Thordike, Freud, Spearman, Pearson....

Desarrollos del Psicodiagnóstico desde 1910	• Hitos significativos: Desarrollo de diversos tests: ✓ Escala de inteligencia de Stanford-Binet ✓ Test Army Alpha y Beta ✓ Evaluación del desarrollo de Gessell ✓ Test de las Manchas de Tinta de Rorschach ✓ Test de la Figura Humana de Goodenough ✓ Test sociométrico de Moreno ✓ Test de Apercepción Temática de Murray ✓ Escalas de Inteligencia de Wechsler ✓ Inventario Multifásico de Personalidad de Minnesota de Hathaway y Mckinley ✓ Cuestionario de Personalidad 16 PF de Catell ✓ Escalas de Clima social de Moos... • Crisis del Psicodiagnóstico Tradicional y desarrollo de la Evaluación Conductual. • Crisis de la Evaluación Conductual • Inicio de un modelo cognitivo de evaluación del funcionamiento intelectual.

Estructuración de Pelechano (1988)	
Periodo de la segunda mitad del S. XIX a 1921 (Antecedentes científicos)	• Sensibilidad romántica a lo individual. • Progresos en el diagnóstico neurológico y psiquiátrico • Evolucionismo biológico y darwinismo social. • Constitucionalismo de corte biologicista. • Aportaciones de la Psicofísica. • El legado del Asociacionismo filosófico del empirismo inglés. • Interés creciente en educación.
Periodo de 1921 a 1945	• Formulación del modelo psicométrico básico. • Potenciación del uso de los tests y pruebas psicológicas. • Generación de una alternativa cualitativa de análisis del caso individual con pruebas proyectivas. • Dos polémicas centrales: estructura factorial de la inteligencia y psicometrismo-proyectivismo. • Gran producción bibliográfica de tests y técnicas proyectivas.
Periodo de 1946 a 1965: Los comienzos de una crisis	• Predominio de EEUU como consecuencia de la emigración de científicos europeos debida a la Guerra. • Polémica en torno al uso social de los tests y a la entidad científica de las pruebas proyectivas. • Aparición de la Psicología Clínica. • Aparición de la Terapia de Conducta con el correspondiente modelo de evaluación. • Cambios en criterios de valoración del sistema educativo y del sistema industrial.
Periodo de 1966 a 1975: La década imperial conductista	• Expansión del análisis funcional o análisis de la conducta • En el modelo tradicional refundición y búsqueda de nuevas opciones. • Carta de naturaleza del situacionismo y pragmatismo. • Comienzan las disensiones en la modificación de conducta produciéndose separaciones significativas. • Reconversión de los modificadores de conducta individual en intervencionistas comunitarios. • Se pierde purismo en los planteamientos teóricos

El final de los 70 y la búsqueda de orden y sentido	• Momento de crisis. • Valoración de programas. • Menor rigidez metodológica y doctrinal. • Incorporación de técnicas tradicionales. • Creación de nuevos instrumentos con diferentes objetivos de análisis. • Desarrollo de la Ecopsicología y la Psicología ambiental. • Críticas a la adecuación de los criterios de eficacia en la evaluación.

CRISIS DE LA EVALUACIÓN TRADICIONAL, INDICADORES DEL RESURGIMIENTO	
Crisis (1965-1980)	• Críticas a las técnicas proyectivas. • Auge progresivo de la Evaluación Conductual. • Ataques sociales a los tests.
Resurgimiento (1980-continua)	• Pervivencia de una amplia utilización de tests en distintos campos. • Incremento del tiempo dedicado a la evaluación psicológica en los congresos. • Aparición de nuevas publicaciones. • Revisiones de tests tradicionales y construcción de nuevos tests. • Disminución del antagonismo evaluación tradicional-conductual. • Aplicación de las teorías del procesamiento de la información a la evaluación de la inteligencia. • Desarrollo de instrumentos de evaluación psiconeurológica. • Ampliación del objeto de estudio de la evaluación psicológica: parejas, grupos, organizaciones, programas.... • Creación de nuevos campos (psicología de la salud...). • Introducción de los ordenadores.

ÚLTIMOS DESARROLLOS
• Continuado desarrollo de técnicas focales. • Creciente sofisticación psicométrica. • Actualización continua de instrumentos de evaluación (MMPI, MCMI, WAIS...) • Constante actualización del DSM, actualmente DSM-IV-TR. • Mayor desarrollo de técnicas administradas por ordenador y mayor sofisticación de la interpretación por ordenador. • Mayor integración entre teoría y práctica. • Más énfasis en la formación multicultural. • Mayor interés por la psicología de la salud y desarrollo de instrumentos de evaluación en este ámbito. • Desarrollos importantes en el campo de la evaluación de programas. • Estudio de las distorsiones de respuesta, y sobre las traducciones e importación sin adaptación de los tests psicológicos. • Nuevos ámbitos de evaluación desarrollados a partir de nuevas tecnologías como lo es la realidad virtual y su aplicación en multitud de áreas de la evaluación desde la capacidad espacial a los tests situacionales. • Consolidación de escuelas y modos de distintos de pensar la evaluación, en donde técnicas proyectivas, tests, análisis de conducta... se complementan con valoración de programas de intervención de índole distinta.

CAPITULO 3.
MODELOS EN EVALUACIÓN PSICOLÓGICA

- 3.1. Introducción
- 3.2. El Modelo Correlacional, Psicométrico o del Atributo
- 3.3. El Modelo Clínico-Dinámico
 - 3.3.1. Perspectiva médico-psiquiátrica
 - 3.3.2. Perspectiva psicoanalítica
- 3.4. El Modelo Fenomenológico, Humanista, Existencial, Constructivista.
- 3.5. El Modelo Conductual
 - 3.5.1. Enfoque conductual radical. Conductismo de primera generación
 - 3.5.2. Enfoque mediacional o de impacto mediatizado. Conductismo de segunda generación
 - 3.5.3. Enfoques integradores e interactivos. Conductismo de tercera generación
 - 3.5.3.1. El modelo Conductual-Cognitivo
 - 3.5.3.2. El modelo Conductual Paradigmático
- 3.6. El Modelo Cognitivo
 - 3.6.1. La evaluación desde el procesamiento de la información
 - 3.6.2. La evaluación neuropsicológica
- 3.7. Resumen: Características diferenciales de los modelos en la evaluación psicológica

3.1. INTRODUCCIÓN

La revisión histórica expuesta previamente nos muestra que han existido distintos enfoques, o modos de enfrentar la tarea diagnóstica. En ocasiones han surgido de concepciones teóricas y/o aplicadas de otras ramas científicas del saber. Otras veces, nuevas investigaciones, han generado cambios y perspectivas conceptuales diferentes. También desde los instrumentos de evaluación se han enriquecido y/o modificado los presupuestos teóricos y confirmado o rechazado la aplicabilidad de los mismos.

En la misma dirección que la señalada por Strauman (1994) podemos considerar que la Psicología como ciencia ha avanzado en parte debido a esta diversidad de perspectivas, y a sus diferentes aportaciones. Se han realizado diversas clasificaciones de los modelos existentes en Evaluación Psicológica. Una de las clasificaciones más exhaustivas de los modelos en Evaluación (Forns, Kirchner y Torres, 1990, 1991; Kirchner, Torres y Forns, 1998) consideran modelos a gran parte de los paradigmas procedentes de la Psicología, estructurando la siguiente clasificación: 1) Evaluación Centrada en las Variables de Persona; 2) Evaluación Centrada en las Variables de Situación; 3) Evaluación Centrada en las variables de Persona y Situación; y 4) Evaluación desde la Psicología Cognitiva. Estas autoras clarifican de cada modelo sus orígenes, los presupuestos conceptuales básicos, las unidades de análisis, aspectos metodológicos, y ámbitos de aplicación de los mismos.

Esta clasificación ha sido criticada por Fernández-Ballesteros (1992a) ya que en su opinión, en esta clasificación se considera modelos de Evaluación Psicológica prácticamente a todos los paradigmas procedentes de la Psicología, pero la mayor parte de los sistemas psicológicos no han llevado consigo su correspondiente versión evaluativa. Desde su punto de vista, para considerar que un modelo o teoría psicológica tiene su correspondiente versión evaluativa tiene que haber producido desarrollos innovadores en la forma de hacer evaluación; es decir, en qué (variables), cómo (métodos), con qué (técnicas) o dónde (ámbito de aplicación). Con esta perspectiva, en su última formulación (Fernández-Ballesteros, 2004a), caracteriza los siguientes modelos: Atributo, Dinámico, Médico, Constructivista, Conductual (Radical / Cognitivo-Coductual) y Cognitivo, clarificando la formulación teórica de cada uno de ellos, las variables que evalúan, los métodos y las técnicas que emplean, sus niveles de inferencia, los objetivos y ámbitos de aplicación.

No obstante, en el contexto de este debate considero que la descripción de los modelos que realizan Forns, Kirchner y Torres es clarificadora, exhaustiva y al menos parcialmente estos modelos han influido de distinto modo en los progresivos cambios conceptuales y metodológicos operados en la evaluación psicológica.

Desde otra perspectiva Avila-Espada (1992a) realiza una revisión crítica de las principales características de los modelos de la evaluación clínica, que aparecen vigentes y suficientemente independientes entre sí durante el período 1970-1985, distinguiendo 5 modelos: 1) El modelo correlacional, psicométrico o diferencial; 2) El modelo clínico-dinámico y su perspectiva psicoanalítica; 3) El modelo fenomenológico, humanista o existencial; 4) El modelo conductual o del aprendizaje social y sus perspectivas, entre ellas la del enfoque a priori; y 5) La perspectiva clínico-experimental derivada de los planteamientos metodológicos que caracterizaron típicamente a la psicología rusa hasta mediados de los años 80. Este último enfoque fue también recogido por Fernández-Ballesteros en la caracterización de los modelos realizada en 1983, sin embargo, fue eliminado en la revisión que realiza en 1992.

Un desarrollo amplio de las características de estos modelos se ha llevado a cabo en trabajos previos (Forns et al., 1990, 1991; Garaigordobil, 1998; Kirchner et al., 1998; Fernández-Ballesteros, 1992a, 2004a). Ante esta diversidad y en aras de ofrecer una visión clarificadora de las características de los modelos más relevantes y que mayor tecnología evaluativa han aportado a este campo, en este capítulo se sintetizan algunas de las ideas básicas en torno a los mismos tales como: 1) orígenes del modelo; 2) presupuestos teórico-conceptuales básicos; 3) objetivos de evaluación y unidades de análisis; 4) aspectos metodológicos y técnicas empleadas; y 5) ámbitos de aplicación, en relación a los modelos que más atención han recibido en la literatura, los más aceptados y con cierta vigencia en el campo de la evaluación. Con esta finalidad se presentan los siguientes modelos:

1. El modelo Correlacional, Psicométrico o del Atributo
2. El modelo Clínico - Dinámico:
 - Perspectiva Médico - Psiquiátrica
 - Perspectiva Psicoanalítica.
3. El Modelo Fenomenológico, Humanista, Existencial - Constructivista
4. El Modelo Conductual:
 - Enfoque Conductual-Radical
 - Enfoque Mediacional o de Impacto Mediatizado
 - Enfoques Integradores.
5. El modelo Cognitivo:
 - La Evaluación desde el Procesamiento de la Información
 - La Evaluación Neuropsicológica

3.2. EL MODELO PSICOMÉTRICO, CORRELACIONAL O DEL ATRIBUTO

Orígenes del Modelo

El modelo de atributos, también denominado psicométrico o de rasgos tiene su origen en los trabajos pioneros de *Galton*, *McKeen Catell* y *Binet*. La obra de Galton, de clara raíz darwinista, destaca por sus aportaciones en la cuantificación de variables antropológicas, que relacionó con unidades de conducta de orden sensorio-motriz (agudeza visual, auditiva, rapidez de reacción, etc.); también son remarcables sus estudios correlacionales y la elaboración de árboles genéticos con el objetivo de analizar las variables hereditarias. McKeen Catell continuó sus trabajos sobre la medida de tiempos de reacción que le permitieron establecer diferencias individuales entre sujetos. Por su parte, Binet aportó a la psicometría algunos conceptos clave. Por ejemplo, defendió un concepto funcionalista de inteligencia, considerándola como entidad compleja compuesta de varias funciones; intentó su medida a través de la selección de diversas conductas graduadas en dificultad genética, de lo cual derivó el concepto “edad mental”.

El objetivo común de estos investigadores se centró en la cuantificación de los aspectos psicológicos del sujeto a fin de resaltar las diferencias individuales, identificar sus causas y determinar cómo se distribuyen estos rasgos entre la población. Como señalan Forns et al. (1990, 1991) es un modelo fuertemente vinculado al desarrollo de la psicometría y a sus formulaciones teóricas.

A partir de estos inicios y hasta los años 60 fue el modelo predominante en evaluación psicológica y cuenta entre sus filas con obras tan relevantes como las de R.B. Catell (1965, 1980), Eysenck (1960, 1967) y Guilford (1967). En la actualidad este modelo sigue en vigencia y es defendido por autores como Kline (1985), o Spielberger y col (1970/1988, 1978). Se puede considerar en esta línea a Exner (1978) porque logra aislar, a partir de sus estudios con el test de Rorschach, tres tipos vivenciales con un fuerte componente intrapsíquico y hereditario que da consistencia e individualidad a la conducta.

En oposición al modelo conductual, durante la década de los 60 y 70 se ha denominado a este modelo “tradicional”, incluyendo también en esta acepción al modelo psicodinámico. El modelo psicométrico ha sido ampliamente difundido, y está vinculado al Psicodiagnóstico desde sus inicios. Su interés se ha centrado en medir los rasgos psicológicos, siendo foco de atención especial la evaluación de la inteligencia.

Presupuestos teórico-conceptuales básicos

La conducta se explica en función de variables internas, personales, orgánicas o también denominadas variables intrapsíquicas. Tales constructos, pueden denominarse como capacidades, atributos, rasgos de personalidad, factores de carácter..., explican y determinan la conducta del sujeto. La conducta se considera como una manifestación indirecta de estos constructos internos que no son medibles de forma directa. Las respuestas del sujeto se interpretan como signos de la variable que se intenta analizar.

Estos constructos internos o variables del organismo dan consistencia y estabilidad al comportamiento a lo largo del tiempo y de las distintas situaciones, al otorgárseles un status estructural. El núcleo base de esta estructura es el concepto de rasgo que tiene un sentido innato, geneticista que propicia esta concepción estable e idiosincrática de la conducta. No obstante existen diferencias en la concepción del rasgo que tienen diversos representantes de este modelo como R.B. Catell o Eysenck.

Estos rasgos son consistentes y se pueden expresar dentro de un “continuum” del rasgo o atributo (mediante las puntuaciones en los test) y son por lo tanto mensurables. Además, se considera que las conductas o comportamientos de los sujetos son susceptibles de comparación entre sí y en relación a un grupo normativo de referencia, sustentando de este modo la base de las diferencias individuales.

Este modelo, no obstante, ha experimentado una serie de reformulaciones que en términos generales pueden concretarse en conceder progresivamente un mayor peso a las variables ambientales. Catell (1980) presenta el modelo PCER (*person-centered environmental representation*) cuya característica es la integración de las variables de persona y de situación. Esta integración, según algunos autores, sitúa al modelo de rasgos en medio de las modernas teorías interaccionistas de la personalidad.

Objetivos de evaluación y unidades de análisis

La finalidad de la evaluación desde este enfoque es encontrar diferencias individuales en aspectos consistentes del comportamiento. Con esta finalidad, los objetivos básicos del modelo son: Descripción, Clasificación-Comparación, y Predicción.

- 1) La descripción consiste en dar cuenta de las características más sobresalientes del comportamiento de un sujeto a partir de los datos obtenidos a través de las diferentes técnicas de análisis, es decir, se miden-describen distintas variables consistentes del comportamiento. Para Forns (1986), la descripción en este modelo equivale a una evaluación sumativa, en el sentido de que da cuenta del status actual del sujeto evitando la interpretación y la explicación de los mecanismos que le han llevado a tal estado.
- 2) La clasificación-comparación de los sujetos en base a sus actitudes y estilos comportamentales, implica la organización de la información en función de criterios previamente definidos.
- 3) La predicción probabilística del comportamiento en base a una metodología de enfoque correlacional. Predecir consiste en establecer una previsión probabilística de una posible conducta futura a partir de una conducta actual. La predicción se basa en la relación entre la puntuación en el test y la conducta criterio que se pretenda predecir, y parte de la tabla normativa que clasifica a los sujetos en el "continuum" de la variable comportamental (rasgo) medida.

Silva (1982) subrayó que el objetivo básico del modelo es la predicción, y que esta predicción supone determinar empíricamente las regularidades que pueden existir entre datos específicos de tests psicológicos y otra característica social, clínica o teóricamente significativa de la persona examinada, especificada con igual claridad.

La evaluación psicológica, desde este modelo, pretende "captar" los rasgos que conforman la personalidad. Un rasgo aislado sólo adquiere su máximo sentido interpretativo cuando es analizado en relación al conjunto de rasgos que determinan una estructura; de lo que se deduce que la evaluación del perfil de personalidad, como conjunto complejo de rasgos es prioritaria frente a la detección del rasgo aislado. El rasgo debe ser analizado en base a la intensidad en su manifestación en relación a un grupo normativo.

La evaluación psicológica debe poner de manifiesto la estructura global de la personalidad para lo que deberá centrar su análisis en las distintas categorías de rasgos que la componen (temperamentales, cognitivos, dinámicos...). Estas unidades de análisis son abordadas en ausencia de cualquier perspectiva evolutiva, representando un corte transversal en la historia comportamental del sujeto. Las respuestas del sujeto en las distintas técnicas de evaluación toman un valor de signo de la existencia de los constructos intrapsíquicos expresados en cada rasgo.

Desde el modelo del atributo se analizan variables intrapsíquicas obtenidas mediante procedimiento empíricos, factoriales o racionales, variables intrapsíquicas o atributos, ya sean éstas rasgos de personalidad, habilidades intelectuales o intereses. Mediante el análisis de los rasgos, dimensiones o factores se pretende un análisis molecular y objetivo de la personalidad del sujeto.

Aspectos metodológicos y técnicas empleadas

La metodología empleada preferentemente por el modelo de los atributos es la correlacional frente a la experimental utilizada desde la perspectiva conductual. El foco de interés se centra en determinar y estudiar las diferencias individuales y en ubicar a un determinado sujeto en relación al grupo normativo en un rasgo o dimensión concreta, que

puede cuantificarse. El método estadístico más utilizado es el análisis factorial en sus distintas variantes.

El objetivo que pretenden es ubicar al sujeto en relación a su grupo de referencia normativo, por lo que el enfoque es preferente nomotético, es decir, consistente en determinar las leyes generales de la conducta, frente al idiográfico centrado en el análisis del sujeto como unidad funcional, y cuantitativo.

Una de las características tecnológicas de este modelo es la aportación de abundante material estandarizado destinado a la evaluación de constructos diversos, tanto de personalidad como cognitivos o motivacionales. Entre las técnicas ya clásicas que han gozado de máxima difusión y aplicabilidad en el área de la personalidad: el Cuestionario factorial de personalidad 16 PF (Cattell 1972); el Eysenck Personality Inventory (EPI) (Eysenck y Eysenck, 1964/1994) con aplicabilidad a diferentes tipos de población (juvenil, adulta) y de diferentes características socio-económicas y culturales; el Minnesota Multiphasic Personality Inventory (MMPI) (Hathaway y McKinley 1942). En el área cognitiva cabe destacar las Escalas Wechsler (WAIS, WISC, WPPSSI) (Wechsler, 1949, 1955, 1963); Primary Mental Abilities (PMA) (Thurstone, 1938/1976); Differential Aptitude Test (DAT) (Bennett, Seashore, Wesman, 1966/1976); Culture Fair Intelligence Test (Test de factor g de Cattell) (Cattell y Cattell, 1959); Standard Progressives Matrices (Raven, 1938/1995) entre otros.

Los niveles de inferencia atribuibles al modelo del rasgo son preferentemente de tipo III (Ballesteros, 1992a, 2004a; Forns et al., 1990, 1991; Kitcher et al, 1998) que implica interpretar la conducta como signo de unos constructos hipotéticos que la determinan. Este nivel de inferencia presupone que la conducta ha sido interpretada en un primer momento como muestra del comportamiento que se pretende evaluar (tipo I) (las puntuaciones de los tests pondrían de manifiesto el patrón de conducta habitual del sujeto); y en segundo lugar esta información se generaliza a categorías comportamentales más amplias (nivel de inferencia tipo II) que se configuran como descripciones taxonómicas.

A partir de este enfoque se construyeron y aplicaron una gran profusión de tests (personalidad, actitudes, aptitudes...) recogiendo información de numerosas variables comportamentales a través de la metodología propuesta con el fin de estandarizar, cuantificar y realizar medidas objetivas (válidas y fiables).

Las técnicas empleadas por este modelo, denominadas psicométricas, son construidas mediante procedimientos estadísticos y factoriales. Básicamente pretenden analizar la conducta de forma indirecta mediante tests, cuestionarios y escalas dotados de consistencia interna, validez de criterio y de contenido, de la fiabilidad, objetividad en su aplicación y valoración. Esta pretensión metodológica-instrumental comporta muestras extensas en la estandarización-baremación a fin de obtener criterios normativos-estadísticos, eludiendo otras encarnizadas polémicas en relación a la normalidad (Maganto, 1995b).

Ámbitos de aplicación

Los ámbitos de aplicación más tradicionales de este modelo han sido el escolar y el industrial con funciones de selección y evaluación de personal, consejo profesional y educacional, evaluación de rendimientos, de climas de trabajo...

También es frecuente trabajar desde esta perspectiva del rasgo en el ámbito clínico con la finalidad de establecer diagnósticos diferenciales y para controlar el proceso de tratamiento. Así

mismo, se han utilizado métodos y técnicas propios de este paradigma en trabajos de investigación, ya que las técnicas psicométricas clásicas son a menudo usadas para definir los parámetros de los grupos o los sujetos participantes en diseños experimentales, con valor de variables independientes.

No cabe duda de que la evaluación desde la perspectiva del rasgo ha tenido momentos históricos de gran resonancia y, de hecho, la sigue teniendo, si bien como indica Forns (1986) el giro progresivo que ha sufrido la evaluación psicológica desde una evaluación de tipo sumativo, que hace énfasis en el estado, a una evaluación de tipo formativo que hace énfasis en el proceso, ha conllevado la propia crisis del enfoque psicométrico que abandonó temporalmente los tests normativos en pro de los tests criterioles.

No obstante, esta supuesta crisis del modelo del atributo es más conceptual que operativa, o sea, más académica que extraacadémica. Pocas son las empresas de selección de personal, pocas las escuelas y escasos los ámbitos clínicos que no utilizan este modelo, tanto desde el punto de vista conceptual, como técnico-metodológico, para funciones clasificatorias, de asesoramiento predictivo y de investigación.

3.3. EL MODELO CLÍNICO-DINÁMICO

Dentro de este modelo se pueden identificar dos enfoques con características diferenciadas: La aproximación clínico-dinámica surgida desde el contexto médico-psiquiátrico y la perspectiva psicoanalítica (Avila-Espada, 1992a, Fernández-Ballesteros, 1992a; 2004a).

3.3.1. LA PERSPECTIVA MÉDICO-PSIQUIÁTRICA

Orígenes del modelo

El modelo nace en el seno de la psicología clínica vigente entre la primera y la segunda guerra mundial, y cobra un rápido desarrollo a consecuencia del masivo incremento en las necesidades de evaluación clínica por efecto de ambas guerras. En un contexto eminentemente médico-psiquiátrico, la psicología clínica adquiere muchos de sus procedimientos en confrontación y dependencia del modelo de intervención de la medicina, en el cual el diagnóstico es un paso necesario, preceptivo, para la clasificación de los pacientes en base a trastornos, y la asignación de tratamiento. Son diversas las circunstancias que inciden en que se formalice el modelo. Por un lado, la vigencia del modelo médico de enfermedad extrapolado a las alteraciones psíquicas. Así mismo, se produce una creciente importancia en los años 40 y 50 de las nosologías psiquiátricas, con la aparición del DSM-I y las clasificaciones de la OMS.

Presupuestos teóricos-conceptuales básicos

Desde este modelo se entiende el comportamiento manifiesto, observable, como signo; los síntomas observables y los datos de los tests se consideran como indicadores que permiten caracterizar los procesos psicopatológicos subyacentes. Los determinantes de la conducta son de naturaleza intrapsíquica, es decir, desde este enfoque se tiende a considerar más las causas genético-biológicas o intrapsíquicas, que las socioambientales. El modelo médico explica la conducta anormal en base a factores endógenos o internos, bien biológicos, bien intrapsíquicos. Es decir, la conducta es una función de condiciones biológicas o personales.

Objetivos de evaluación y unidades de análisis

El diagnóstico se opera con la finalidad de clasificar en función del trastorno, para aplicar posteriormente un tratamiento, teniendo por objetivos: (1) Conocer el grado de afectación del sujeto en un trastorno; (2) Detectar factores etiológicos (histórico-biográficos); (3) Diferenciar trastornos orgánicos y trastornos funcionales; (4) Determinar el curso que ha seguido el proceso psicopatológico; y (5) Intervenir o tratar con el tratamiento adecuado para ese proceso psicopatológico.

En el caso del modelo médico, la explicación de la conducta anormal, al menos teóricamente, daría lugar a un diagnóstico en el sentido de que éste ha de estar basado en una etiología específica, ayudando éste al pronóstico (o predicción) y a su control. Ahora bien, dado que en este modelo el psicólogo actúa, fundamentalmente, como un auxiliar del médico, en el sentido de ser su contrastador, el objetivo se reduce al de la pura descripción y clasificación del sujeto. Es tan sólo en un modelo médico psiconeurológico cuando la evaluación del comportamiento anormal, tomado como expresión de una lesión del SNC, puede conllevar su explicación y rehabilitación.

A pesar de que el modelo señala como objetivo la obtención de un diagnóstico para la intervención o tratamiento, durante muchos años se abre una grieta entre la labor diagnóstica y su aplicación a la terapia. El modelo médico implica el estudio en cada sujeto que presenta trastornos de conducta y cómo puede ser clasificado según una serie de entidades nosológicas consideradas como enfermedades mentales (esquizofrenia, depresión, demencia...) y/o en qué medida presenta alteraciones biológicas que se supone subyacen a dichos trastornos (enfermedad de Alzheimer, síndrome de Down...). Las prácticas diagnósticas que se derivan de este enfoque se ocupan de la identificación de signos y síntomas que sean indicadores de los procesos patológicos subyacentes, procediendo bien por vía hipotético-deductiva o inductiva a la obtención del diagnóstico, extraído por referencia a un modelo derivado de las nosologías psicopatológicas.

Aspectos metodológicos y técnicas utilizadas

Desde el modelo médico se emplean fundamentalmente tests y técnicas de recogida de información para la contrastación de las hipótesis de partida; se tienden a emplear procedimientos tipificados contruidos mediante criterios racionales, empíricos o factoriales en la medición de rasgos, factores, dimensiones o entidades nosológicas. A través de las pruebas se pretende conocer la posición que tiene un sujeto respecto a un grupo de referencia normativo, en el atributo objeto de examen.

Los métodos y las técnicas de evaluación utilizadas son múltiples, tanto de tipo extensivo como intensivo, observacionales, psicométricas, proyectivas o subjetivas, en base a cuyos datos se realiza el diagnóstico. Buena parte de las críticas que en los años sesenta se dirigen contra las técnicas proyectivas (los instrumentos de evaluación más característicos de este modelo) se derivan de las implicaciones de su uso en un contexto medicalizado. Las respuestas de los sujetos ante las pruebas son valoradas asituacionalmente, sin tomar en consideración el valor estimular específico de las instrucciones y materiales de administración, ni otros requisitos básicos de naturaleza psicométrica (adecuación a la cultura, validez ecológica, ausencia de normas de respuesta, sesgos y tendencias de respuesta...) necesarios para establecer

cualquier hipótesis. Cuando la interpretación se apoya en tales protocolos se corre el riesgo de ser extremadamente superficial o necesitar de un nivel de inferencia no justificado por los datos. La reformulación habida de las técnicas proyectivas a partir de los años sesenta las aleja de las premisas clásicas de este modelo, contribuyendo a la revisión del mismo.

El modelo teórico desde el que se trabaje va a influir en las técnicas seleccionadas para la recogida de información sobre el sujeto, y en el nivel de inferencia que se efectúe sobre las conductas registradas. El modelo médico, al igual que el del atributo utilizan predominantemente inferencias de nivel III, es decir, se utilizan términos con la consideración de constructo hipotético o estado interno, el cual implica una etiología o condición causal con base intrapsíquica. Así, las conductas del sujeto son consideradas como signos de la existencia de un atributo subyacente en el mismo.

Ámbitos de aplicación

El modelo médico que, en parte, es una derivación clínica del modelo del atributo, se ha dedicado exclusivamente al ámbito clínico.

3.3.2. LA PERSPECTIVA PSICOANALÍTICA

Orígenes del modelo

El Psicodiagnóstico desde la perspectiva psicoanalítica, tiene su origen en los presupuestos teóricos sustentados por el modelo freudiano de personalidad. El origen de este enfoque se encuentra en el campo de la psicopatología, especialmente en los trabajos pioneros de Charcot y Janet. Pero será Freud quién llegará a construir un corpus teórico con entidad propia, el modelo psicoanalítico, que tendrá importantes resonancias en el campo aplicado. En 1924, el propio autor en su Autobiografía (1981) señala dos etapas bien diferenciadas en el desarrollo de sus teorías. De 1895 a 1907 elabora los conceptos inconsciente, preconsciente y consciente desarrollados en sus obras “Estudios sobre la Histeria” que data de 1895 y “La interpretación de los sueños” publicada en 1899. De 1907 en adelante elabora los conceptos de yo, super-yo, y ello, explicitados en sus textos “Introducción al narcisismo” de 1914 y “Más allá del principio del placer” de 1920.

Este enfoque surge en un momento en que el auge se centraba en el descubrimiento de las causas orgánicas de las enfermedades mentales, denominada organogénesis. La innovación de Freud consistió en defender que otras fuentes psíquicas, emocionales y afectivas provenientes del propio individuo, eran los fenómenos causales de las enfermedades mentales. La psicogénesis se desarrolla en gran manera a partir de él. Los elementos que amplían la comprensión de las perturbaciones psíquicas radican en el descubrimiento del inconsciente, la noción de conflicto intrapsíquico patógeno y el concepto de mecanismo de defensa. El modelo psicoanalítico actualmente abarca numerosas variaciones y diversas perspectivas. No obstante, nos referiremos preferentemente al modelo expuesto por Freud, sin hacer referencias a otros eminentes psicoanalistas como A. Freud, M. Klein, E. Erikson, D. Winnicott o J. Lacan.

Presupuestos teórico-conceptuales básicos

En el contexto de esta perspectiva coexisten dos posturas respecto al Psicodiagnóstico, por un lado, quienes consideran que el diagnóstico es un obstáculo para el psicoanálisis, o quienes creen que un cierto diagnóstico es necesario y puede ayudar para orientar el proceso psicoanalítico. Los presupuestos básicos de este enfoque más frecuentemente señalados son:

- 1) La conducta es considerada como síntoma, es una manifestación externa de un conflicto intrapsíquico que se origina por la pugna entre las tres instancias psíquicas, dotadas de energía, Ello, Yo, y Superyo, que actúan según diferentes procesos, inconsciente, preconscious y consciente (Cuadros 3 y 4).

Cuadro 3. Organización topográfica de la personalidad

Organización topográfica de la personalidad: Inconsciente, Preconscious y Consciente	
Inconsciente	En el sentido descriptivo, califica cualquier contenido no presente en el campo actual de la conciencia. En un sentido tóxico, es un sistema constituido por los contenidos reprimidos a los que ha sido negado el acceso al sistema preconscious y consciente por la acción de la represión.
Preconscious	Se refiere a las operaciones y contenidos que no están presentes en el campo actual de la conciencia, sin embargo, son accesibles a la misma cuando se levanta la censura.
Consciente	Hace referencia a los contenidos (pensamientos, recuerdos...) de los que tenemos conciencia en un momento determinado. El sujeto se da cuenta o es consciente de una parte de sus procesos mentales en el momento presente.

Cuadro 4. Características de las Instancias Psíquicas: Ello, Yo y Super-Yo

Características de las Instancias Psíquicas: Ello, Yo y Super-Yo	
Ello	Es fuente de toda energía y sede en donde radican los instintos. En el principio de la vida la persona está configurada por el Ello, energía que tiende a descargarse y a buscar satisfacción. La búsqueda de placer se denomina Principio de Placer. Este es su principio de funcionamiento básico, ya que si no se consigue la satisfacción inmediata deseada no se descarga la energía y queda un remanente de energía que produce malestar en el sujeto y provoca conflicto. El Ello al ser de naturaleza instintiva, no se rige por la razón, la lógica y el orden. Es de naturaleza inconsciente y su contacto con la realidad se vehiculiza a través del Yo.
Yo	Ejerce su actividad en conformidad y relación al Principio de Realidad. Atempora y/o modula las demandas instintivas del Ello en función del contexto social, procurando dar satisfacción a ambas fuentes de exigencia: pulsionales y sociales. Se va desarrollando evolutivamente y madurando a través de la interacción con el mundo externo. Los complejos procesos psicológicos cognitivos, como la atención, la memoria, percepción... son propios del Yo y ejercen de mediadores entre el Ello y el mundo exterior.
Superyo	Es la instancia que, obteniendo la energía del Ello, se enraiza en el Yo. Esta energía de naturaleza biológica configura dos fuerzas instintivas preferentemente: El instinto de vida y el instinto de muerte. El superyo viene a ser para Freud el código moral desarrollado por la internalización e identificación de figuras y normas parentales y sociales, que se reciben a lo largo de la historia individual, y de donde derivan el sentido de la censura-culpa o el sentido de satisfacción

- 2) La conducta es fruto de las pulsiones internas que han determinado la estructura de la personalidad. Esta se ha ido gestando a lo largo del desarrollo infantil priorizando una determinada zona erógena.
- 3) El conflicto intrapsíquico se origina en la lucha entre impulsos inconscientes y deseos conscientes, es decir, entre las demandas provenientes de uno mismo y las exigencias sociales.
- 4) Es posible el conflicto entre las instancias psíquicas, y en base al tipo de conflicto y a los mecanismos de defensa que propicien, Freud explica la estructura y funcionamiento de los trastornos psíquicos.
- 5) El desarrollo de la personalidad es visto desde un enfoque evolutivo que no siempre ha sido concebido igualmente por los autores psicoanalíticos. Freud concreta una serie de estadios del desarrollo psicoafectivo, ligados fuertemente a las necesidades biológicas primarias del individuo y en sus inicios, a las distintas zonas erógenas que las hacen factibles. Desde su punto de vista el desarrollo pasa por etapas sucesivas ordenadas cronológicamente (Cuadro 5).

Cuadro 5. Estadios del desarrollo psicoafectivo

Estadios del Desarrollo Psicoafectivo: Oral, Anal, Fálica y Latencia	
Etapas oral	Representa el primer año de vida. Las gratificaciones provenientes de la boca simbolizan además el tipo de relación de objeto que el bebé establece con el mundo exterior, casi exclusivamente la madre. La actividad prioritaria de succión conlleva la incorporación.
Etapas anal	Se opera en el segundo año de la vida. Las actividades prevalentes derivan de actividades de expulsar o retener en respuesta a la educación esfinteriana propia del momento. La relación objetal se establece en estrecho paralelismo a la relación del niño con la materia fecal y con las exigencias educativas paternas.
Etapas fálica	A los 3 años se inicia la etapa fálica, denominada así por ser la región genital fuente de placer y de tensión. Se evidencia la diferencia de sexos y se establece el Complejo de Edipo. El papel estructurante que tiene en la personalidad es definitivo para comprender cada estructura patológica por sus modalidades de resolución. De la relación con los objetos, figuras parentales, devienen fuertes sentimientos de amor y odio, envidia y rivalidad, identificación y oposición, que configuran una etapa difícil para el niño y en la que las fuerzas represoras juegan un papel decisivo en su finalización.
Etapas de latencia	Cubre desde los 6 años hasta la pubertad. En ella se internalizan las identificaciones psico-socio-sexuales con los miembros del propio sexo, a través de las identificaciones con los padres, y se desarrollan las sublimaciones que permiten las adquisiciones escolares, recreativas, y culturales.
Etapas genital	Se inicia en la pubertad-adolescencia enlazando con la edad adulta. La prevalencia de la gratificación vuelve a la zona genital, reestructurándose la personalidad. La maduración progresiva del desarrollo heterosexual orienta al Yo a buscar nuevos objetos amorosos, diferentes de los padres y a establecer un vínculo libidinal personal.

- 6) El comportamiento puede ser explicado en base a una serie de construcciones teóricas internas que conforman la estructura de la personalidad, las cuales, junto con una serie de dinamismos internos inconscientes, determinan la conducta, por lo que cualquier manifestación conductual será entendida como una expresión de una condición interna de la persona. El comportamiento observable es signo de los procesos inconscientes que lo determinan, y a la vez, muestra de los recursos adaptativos que posee el sujeto. Ambas

clases de información son relevantes para la formulación diagnóstica descriptiva que se haga del sujeto. Así se opera con un modelo para la conducta que puede expresarse como c f (O en A) (O modulado históricamente por A).

Objetivos de evaluación y unidades de análisis

La finalidad de la evaluación en este modelo es el análisis de la estructura intrapsíquica del sujeto, entendida como la resultante del equilibrio dinámico entre sus distintas instancias. Interesa comprobar su adecuación o inadecuación, su fortaleza o debilidad, la interacción de las tres instancias psíquicas (Ello, Yo, Super-Yo), el peso específico de sus elementos topológicos (Consciente, Inconsciente y Preconsciente) y la calidad y eficacia de los mecanismos de defensa con los que cuenta el individuo.

Un punto especial de este enfoque es el énfasis que se pone en la historia pasada del sujeto. La trayectoria infantil de satisfacciones y frustraciones de los impulsos básicos, el papel de las figuras parentales, y las vivencias internas (introyecciones) que de ellas ha efectuado el sujeto, adquieren gran relieve como causas explicativas de la conducta actual. Por otra parte, la evaluación desde la perspectiva psicoanalítica no se basa en el análisis de la conducta manifiesta del sujeto, sino en sus elaboraciones mentales. No interesa tanto la realidad objetiva cuanto la subjetiva, que es la que tiene auténtica significación en esta perspectiva.

Así, el objetivo principal se centra en explicar la etiología del comportamiento humano a partir de la descripción de sus estructuras intrapsíquicas y de su funcionamiento (Forns, et al., 1990). Estos objetivos han sido objeto de críticas. La explicación freudiana se basa en el concepto de determinismo fisicalista de causa-efecto imperante en la época. Hoy en día el concepto se ha reformulado en términos de determinismo estructural y se considera que la estructura de un campo determina la de sus elementos, lo cual no implica que sea su causa. La física moderna, la mecánica cuántica y el principio de indeterminación de Heisenberg han dado lugar a una concepción probabilística de las leyes.

Sus objetos de estudio fundamentales son: (1) Los procesos y contenidos inconscientes, facilitar su evocación para poder efectuar su descripción y establecer su dinámica; (2) La descripción de la estructura psíquica, en particular del Yo, a través de la valoración de su fuerza, y de los recursos adaptativos y defensivos a su servicio; y (3) La valoración de la accesibilidad del sujeto al tratamiento psicoanalítico. Desde el enfoque dinámico se emprende el estudio de la estructura de la personalidad desde una perspectiva molar en base a los distintos conceptos psicodinámicos que se ponen de manifiesto a través del comportamiento manifiesto de sujeto. Las variables objeto de estudio son de naturaleza intrapsíquica.

Aspectos metodológicos y técnicas empleadas

El modelo dinámico ha basado sus supuestos teóricos en el método clínico a partir de observaciones cualitativas (Fernández-Ballesteros, 1992a; 2004a). Emplea el método inductivo ya que a partir de las observaciones clínicas se infieren los elementos dinámicos y estructurales que dan cuerpo a la conducta y que posibilitarán un planteamiento comprensivista de la misma. La investigación psicodinámica, en cuanto a las técnicas de análisis, es básicamente clínica, no obstante, existen también estudios de contrastación empírica que utilizan preferentemente el enfoque correlacional y la validez de constructo.

El nivel de inferencia es de tipo IV, ya que la conducta de un sujeto es interpretada como un signo a partir del cual se sacan conclusiones que están respaldadas y pueden integrarse en una teoría de la personalidad.

La evaluación es tendencialmente idiográfica, haciéndose énfasis en el individuo concreto del que se intentará captar toda su personalidad de un modo global. Se utilizan métodos fundamentalmente intensivos, idiográficos en teoría, en el estudio del comportamiento del sujeto individual (biográficos, recuerdos tempranos, observacionales, tipología del vínculo establecido con el analista...), aunque también pueden emplearse métodos que posean una vertiente nomotética, al menos potencialmente (técnicas proyectivas, cuestionarios...) (Avila-Espada, 1992a).

Las técnicas utilizadas preferencialmente para conseguir el objetivo evaluador son la entrevista libre o semiestructurada, la asociación libre, el análisis de los sueños, de los "lapsus linguae", de los olvidos; en resumen, todo aquello que facilite la introducción en la vida mental inconsciente del sujeto. Freud inicialmente utilizó la hipnosis, para posteriormente sustituirla por el análisis de la transferencia, entendida como un proceso a través del cual se actualizan en la relación analítica los deseos inconscientes que en el sujeto suscita el evaluador o terapeuta, y que son una repetición de sus primeras relaciones de objeto.

Las técnicas proyectivas, instrumentos frecuentes de análisis utilizados por psicólogos adscritos al modelo, son un desarrollo posterior a Freud, pero de gran profusión, y se emplean como vía de acceso al mundo mental inconsciente del sujeto. Existen varios tipos de técnicas que han sido clasificadas en función de la actividad que conllevan: Estructurales o perceptivas (Rorschach); Temáticas en las que se elaboran historias (TAT o TRO); Expresivas ya que se solicita al sujeto dibujar una o varias figuras (DFH. HTP. Test de la Familia); Constructivas porque solicitan al sujeto construir con diversos materiales que se le aportan (Test del Pueblo); y Asociativas debido a que el sujeto presenta sus asociaciones frente a palabras, frases o cuentos (Frases incompletas, Fábulas de Duss...).

Estos instrumentos adquieren valor diagnóstico eficaz cuando se encuentran enmarcados en la perspectiva teórica psicoanalítica. Para Avila-Espada (1992a) en ese modelo las técnicas funcionan como meros reactivos que evocan respuestas de interés para el psicoanalista. En consecuencia no se estudia el problema de la fiabilidad ni de la validez empírica de los procedimientos, otorgándose importancia fundamentalmente a la validez de constructo. Sin embargo, otros autores (Kline, 1992) se preocupan de explicar en sus trabajos el matiz diferente que constructos psicométricos tales como sensibilidad, fiabilidad y validez toman en lo que a las técnicas se refiere.

Ámbitos de aplicación

Por basarse en el estudio de la conducta anormal, este modelo actúa fundamentalmente en situaciones clínicas. Su ámbito de aplicación es eminentemente clínico y tiene por finalidad mejorar la adaptación personal del sujeto mediante la toma de conciencia de sus dificultades y de la etiología de las mismas. Ello conduce a una reestructuración de las diferentes fuerzas e instancias intrapsíquicas que determinan la conducta actual y permite una mejor fórmula expresiva de las pulsiones (Forns et al., 1990, 1991, Kirchner et al., 1998).

El trasvase al campo educacional (A. Freud) y las derivaciones de programas de la Escuela de Summerhill (Neill), así como la elaboración de programas de desarrollo afectivo y la

utilización de este marco teórico en los procesos de orientación y selección, son una muestra de la expansión de este modelo a otros ámbitos.

Pese a este encuadre clínico, algunas de las técnicas proyectivas se han utilizado y se utilizan en contextos de selección y orientación como complemento de los tests psicométricos tradicionales. Este hecho ha sido criticado por algunos autores que ven en ello una postura de inadecuación (Forns, 1986) o de eclecticismo que impide que los campos de actuación se definan (Barrios y Hartman, 1986).

Desde el modelo psicodinámico, se observan repetidos esfuerzos por operativizar las variables relativas al proceso diagnóstico, las categorías diagnósticas en función de la clasificación elegida, así como la planificación y evaluación del propio proceso terapéutico. Una actualización de estos intentos ha sido realizada por diversos autores (Dahlbender, Rudolf, und der AK, 2006; Schneider, Klauer y Freyberger, 2008; Task Force, 2007). En la misma línea, Westen y Thompson-Brenner, (2006), McWilliams (2007) y Etnyre (2008), plantean la complejidad de atender a los procesos internos y externos que se integran en la evaluación psicodinámica y aplicar métodos de investigación que posibiliten comparar los diagnósticos realizados entre categorías del DSM-IV-R y las clasificación psicodinámica. La dialéctica largamente mantenida entre clasificaciones multiaxiales y operativas y clasificaciones desde la perspectiva ideográfica ha sido recientemente actualizada por Halasz (2008) y McCarthy y Hoffman (2008).

3.4. EL MODELO FENOMENOLÓGICO, HUMANISTA, EXISTENCIAL, CONSTRUCTIVISTA

Orígenes del Modelo

Creada por Husserl, la fenomenología y su derivación, el existencialismo, configuró una concepción del hombre como ser personal, singular, esencialmente determinado por su propia percepción del mundo. Una notable influencia tuvo lugar, en Estados Unidos, en las concepciones personalógicas de Murray (1938) y Allport (1939) y en los desarrollos de la Psicología Individual Adleriana. Tuvo también gran influencia en la Psicoterapia Guestáltica de Perls que propugnaba una concepción holística de la conducta, con desarrollos que podrían denominarse pre-cognitivos. En el clima intelectual generado por la Fenomenología, surgieron los planteamientos de Rogers (1942, 1951, 1961, 1970) creador de la *Psicoterapia centrada en el cliente*, así como las propuestas de Kelly (1955, 1971) desde las que se considera que la conducta humana esta determinada por los constructos personales o formas de anticipar el mundo.

Presupuestos teórico-conceptuales básicos

Las aproximaciones fenomenológicas a la conducta defienden que los fenómenos de percepción sobre sí mismo y sobre la realidad condicionan vivamente la conducta humana (Forns, 1990). Entre los autores más representativos del modelo fenomenológico o humanista destacan Rogers, Maslow, Perls y May. Analizar las aportaciones de cada uno sobrepasaría el objetivo de esta breve revisión de los modelos evaluativos, por lo que me centraré en los postulados de Rogers a modo de ilustración.

Desde el modelo fenomenológico se entiende el comportamiento como un producto total, determinado básicamente por la percepción subjetiva del ambiente, los estilos

comunicacionales y el potencial humano del sujeto. No obstante, una de las principales carencias del modelo está en su insuficiente formulación de un concepto de conducta específico al mismo (Avila-Espada, 1992a). Así pues, desde esta perspectiva se considera que cada ser humano tiene una visión distinta y única de la realidad. La conducta de un sujeto está en función de su *campo fenoménico* que se configura a partir del conjunto de experiencias, tanto internas (percepción de lo que ocurre dentro del cuerpo) como externas (percepción de lo que ocurre en el mundo exterior) vivenciadas por el organismo.

El Yo, otra de las nociones clave en la teoría de Rogers, es concebido como una diferenciación progresiva a partir de la experiencia del organismo. En la conformación del Yo es esencial el papel desempeñado por los otros, ya que para desarrollarse plenamente se necesita de la estimación y respeto de los demás. Lo que impulsa la conducta, para Rogers, es la tendencia innata del organismo hacia la autorrealización (fuerza interna que empuja al individuo a desarrollar al máximo su potencial). El proceso de autorrealización puede verse interferido si la valoración de la propia experiencia es relegada en favor de los valores de otros. Cuando la idea que el sujeto tiene de sí mismo se opone con su experiencia, distorsionará la realidad para preservar su autoconcepto. No se acepta el concepto de enfermedad mental, sino de diferentes grados de incongruencia entre el autoconcepto y la experiencia.

Objetivos de evaluación y unidades de análisis

Como señala Forns (1990) en este marco teórico no hay evaluación psicológica ni proceso diagnóstico. El evaluador y el terapeuta se confunden en una única función que tiene más de comprensiva y empática que de evaluadora. No se acepta "medir" las cualidades del sujeto por lo que implica de establecimiento de juicios de valor y de comparación con otras individualidades, ni se acepta categorizarle o asignarle a un grupo determinado, porque, sólo el individuo en cuestión se conoce suficientemente a sí mismo, y porque todo sistema de categorías es inútil y absurdo.

Así, el proceso diagnóstico no tiene como objeto la producción de etiquetas, descripciones o predicciones, criticándose el concepto de enfermedad mental, las categorías psicopatológicas y la idea de rasgos estáticos de personalidad. En esta perspectiva no existe una separación entre la actividad diagnóstica y la actividad terapéutica. Las técnicas utilizadas con y por los sujetos son elegidas, valoradas, e interpretadas por o con los propios sujetos, los cuales derivan su propio crecimiento y aprendizaje de tales actividades. Actividad diagnóstica y actividad terapéutica quedan así imbricadas en una única ejecución. Se rompe, además, con el tradicional rol asimétrico que separa al evaluador del evaluado, en aras de los requisitos psicométricos de la estandarización, que aquí pierden todo su sentido.

Los objetos de estudio más característicos son la autopercepción, el autoconcepto, la percepción de los otros y del ambiente, las estrategias personales de resolución de problemas, la evaluación de la tendencia a la autorrealización, del potencial humano y de la motivación de crecimiento; la integración del sujeto y la reducción de la tensión persona-ambiente, persona-sistema o la resultante de su integración en los diferentes sistemas.

Aspectos metodológicos y técnicas empleadas

Si bien no puede decirse que el modelo fenomenológico haya generado una metodología en el sentido estricto, tal y como la entendemos hoy, es decir, a partir de los supuestos de la

Psicología Científica, en todo caso, en un sentido amplio, sí puede hablarse de metodología en cuanto se determina un objetivo básico: lograr un mejor ajuste personal y social de los sujetos para lo que se establecen unas técnicas adecuadas a este objetivo. Se tiende a que el proceso y las técnicas utilizadas durante el mismo sean valoradas e interpretadas por los propios sujetos a fin de que tal actividad acreciente la toma de conciencia de sí mismo.

La técnica utilizada por Rogers es la inferencial, intuitiva, mediante la cual, a partir de las verbalizaciones que realiza el sujeto, se intenta captar su mundo interno. Dentro de los planteamientos Rogerianos se acentúa la importancia de la intuición y la empatía como factores que crean las condiciones para una evaluación centrada en el cliente.

Las técnicas más características son las que se basan en la *observación* y la *autoobservación*. Otra técnica es la *entrevista libre*, siendo el sujeto el que decide qué, cuando y cómo hablar. El evaluador-terapeuta se limitará a escuchar, servir de reflejo al cliente y potenciar un proceso de introspección. Otro grupo de técnicas empleadas preferentemente por este enfoque son las *subjetivas*, mediante las cuales se explora la percepción que tiene el sujeto de sí mismo y de su ambiente. Las técnicas como la técnica de clasificación Q, el diferencial semántico, el test de los constructos personales, y las listas de adjetivos han sido utilizadas ampliamente. Así mismo se emplean técnicas proyectivas y psicométricas, aunque éstas siempre se utilizan respetando la filosofía, objetivos y procedimientos característicos de este enfoque.

Ámbitos de aplicación

El ámbito fundamental de aplicación del modelo fenomenológico es el clínico. Posteriormente se ha extendido a otros ámbitos como el educativo y el de la psicología comunitaria. En el ámbito educativo intenta establecer las bases para una mejor relación profesor-alumno, en la que el profesor pasa a ser una persona que crea el clima psicológico apropiado para que el alumno desarrolle al máximo sus potencialidades tanto cognitivas como emocionales. En el área de la psicología comunitaria se ha aplicado la teoría de Rogers con la finalidad de promover cambios comunitarios y sociales que permitan un mejor desarrollo de las personas.

3.5. EL MODELO CONDUCTUAL

El modelo conductual se situó inicialmente en una posición opuesta a los modelos existentes, es decir, el modelo psicométrico y el modelo psicodinámico, a los cuales encuadró bajo el término "evaluación tradicional". Atacó su determinismo internalista, la importancia dada a la historia del sujeto, sus niveles de inferencia, sus comparaciones nomotéticas, la estabilidad que atribuyen a la conducta, la separación entre evaluación y terapia, los instrumentos empleados... Antes de introducimos en las características específicas de las distintas evoluciones del modelo conductual, se presenta la diferenciación planteada por Barrios y Hartmann (1986) y recogida por Westmeyer (1992) (Cuadro 6).

**Cuadro 6. Diferencias entre los Enfoques Conductual y Tradicional de Evaluación
(Barrios y Hartman, 1986, citado en Westmeyer, 1992)**

	CONDUCTUAL	TRADICIONAL
Concepción de la personalidad	Los constructos de personalidad se emplean para resumir modelos específicos de conducta.	La personalidad es un reflejo de estados o rasgos subyacentes y duraderos.
Motivos de la conducta	Mantenimiento de las condiciones deseadas en el ambiente habitual o en las interacciones ambiente - organismo.	Intrapsíquicas o internas al individuo.
Papel de la conducta	Importante como un ejemplo del repertorio de la persona en una situación específica.	Asume importancia sólo en tanto signifique causas subyacentes.
Papel de la historia	Relativamente sin importancia excepto, por ejemplo, para dar una línea de base retrospectiva.	Crucial, puesto que las condiciones actuales son vistas como un producto del pasado.
Consistencia de la conducta	Se considera a la conducta como específica de la situación.	Se espera que la conducta sea consistente a través del tiempo y de los ámbitos.
Empleo de datos	Para describir las conductas de interés y las condiciones que las mantienen. Para elegir el tratamiento adecuado. Para evaluar y revisar el tratamiento.	Para describir la etiología y el funcionamiento de la personalidad. Para diagnosticar o clasificar. Para realizar pronósticos. Para predecir.
Nivel de Inferencia	Bajo.	Medio a alto.
Comparaciones	Intraindividuales o idiográficas.	Interindividuales o nomotéticos.
Métodos de evaluación	Métodos directos (por ejemplo, observaciones de la conducta en situaciones naturales).	Métodos indirectos (por ejemplo, entrevistas, autoinformes).
Temporalización de la evaluación	Continua.	Pre y, quizás, postratamiento o estrictamente al diagnosticar.
Objeto de la evaluación	Específico y extensivo (multiambiente, multitiempo, multirespuesta)	Global (p. ej. de cura o mejora) pero sólo del individuo.

Si bien son numerosas las características que diferencian la denominada evaluación conductual de la tradicional, en la actualidad se va perfilando una tendencia que las considera más complementarias que antagónicas. A este respecto me parece interesante la perspectiva de Tous (1989a) que puede concretarse en el cuadro 7.

Cuadro 7. Complementariedad de la Evaluación Tradicional y la Evaluación Conductual (Tous, 1989)

	TRADICIONAL	CONDUCTUAL
Objetivo	Explicitar diferencias interindividuales	Explicitar diferencias intraindividuales
Naturaleza	Ley General	Estados particulares de la Ley General
Información	Definida conceptualmente	Definida operacionalmente
Significado de la Puntuación	Refiere al grupo o muestra normativo/a	Refiere a la condición de producción (para cada individuo). Absoluto
Características de las condiciones	Derivadas de una teoría. Capaces de discriminar tendencias de conducta	Derivadas de la condición empírica. Capaces de discriminar conductas específicas
Optimización	Las tendencias de conducta deben especificarse en conductas concretas. Definición operacional de las leyes generales	Los repertorios de conducta deben definirse operacionalmente en relación a las leyes generales
¿Qué mide?	Una tendencia, una propensión, una vulnerabilidad, un estilo	Una muestra de población de conducta, una muestra de dominio de conductas, una habilidad específica, una respuesta dominante

El enfoque conductual surgió al abrigo del análisis experimental de la conducta con el objetivo de aplicar a la evaluación psicológica los avances teórico-prácticos que se habían logrado en el laboratorio en cuanto a análisis y control de la conducta. Kanfer y Saslow (1965) y Mischel (1968/1973) fueron pioneros en la crítica a los presupuestos métricos de los instrumentos usados en la evaluación tradicional y en la búsqueda de un nuevo proceso de evaluación que aceptara los principios de la teoría del aprendizaje.

La teoría conductista es una teoría compleja que ha tenido en el tiempo sucesivos cambios conceptuales, dando lugar a diversos modelos evaluativos. En líneas generales ha pasado de una concepción radical de la conducta que mantenía como foco de interés prioritario la observación y medición de los comportamientos observables, a la inclusión de variables mediacionales, es decir, se acepta la existencia de variables mediacionales entre el estímulo y la respuesta, y finalmente el modelo evoluciona en la aceptación de variables de tipo cognitivo. No obstante, hay que señalar que las distintas perspectivas que se operan en esta evolución, no anulan las previas, coexistiendo en la actualidad estas posturas teóricas.

En 1913 Watson publica el artículo “*Psychology, as the behaviorist views*” que supone el manifiesto fundacional del Conductismo. Su afán objetivista le llevó a negar la validez de la introspección como método de conocimiento de la conducta humana y a negar conceptos mentalistas (sensación, emoción, idea, deseo...) sustituyéndolos por una explicación de la conducta en términos de conexión estímulo-respuesta. Así, en esta etapa (conductismo de primera generación) se explica la conducta por principios objetivos vinculados al aprendizaje y a la formación de hábitos, y está caracterizada por el experimentalismo con animales.

En la década de los años 30 emerge el conductismo de segunda generación o Neoconductismo. La inclusión en el diseño experimental de las variables intervinientes propició la ruptura con el ateorismo de Watson. Los trabajos de Tolman, Guthrie, Hull coinciden en la

inclusión de variables intervinientes, es decir, elementos de conexión explicativa entre estímulo y respuesta.

Esta corriente explicativa mediacional no fue compartida por toda la escuela conductista. Skinner (1904-1990) se mantuvo en una explicación de la conducta en términos netamente conductuales rehuyendo incluso de explicaciones fisiológicas, representando la defensa de un empirismo radical de tipo descriptivo-funcional. Su principal aportación es el condicionamiento operante que vincula la respuesta del sujeto a la ocurrencia del refuerzo. La conducta se configura por las presiones ambientales que actúan como reforzadores positivos o negativos, y puede ser mantenida y controlada mediante la manipulación de sus consecuencias.

Si bien los representantes del conductismo de segunda generación mantienen unánimemente el uso exclusivo del método experimental para el análisis de la conducta, sin embargo, se observa una dispersión importante respecto a los términos explicativos de la misma.

Esta dispersión provocó, hacia los años 50, una situación de crisis que dio pie a un conductismo de tercera generación desde el que se plantearon variadas propuestas, entre las que se pueden destacar la inclusión de variables cognitivas, o la elaboración de un diseño unificador de explicación de la conducta como el conductismo paradigmático.

En síntesis, se pueden distinguir tres modelos conductuales en función del peso específico que conceden a las variables ambientales como determinantes de la conducta. El primer lugar, el conductual-radical, que considera el ambiente como determinante primordial directo sobre la conducta del sujeto. En segundo lugar, los modelos de impacto mediatizado o mediacionales, que estiman que el impacto de la situación sobre la conducta está mediatizado por toda una serie de variables del organismo. Y en tercer lugar, los modelos interactivos o integradores, que, incluyendo el ambiente como variable elicitadora de conductas, estiman que su impacto sobre el organismo no es directo ni siquiera mediatizado, sino interactuante con las propias variables de cada organismo.

Pese a las distintas evoluciones se podría afirmar que estos diferentes enfoques comparten una serie de características comunes como son: (1) Insistencia en el uso de técnicas objetivas, (2) Defensa de las variables estímulo-respuesta, (3) Aceptación de los principios del condicionamiento, como base de las leyes del aprendizaje; y (4) Énfasis en los determinantes externos de la conducta, y por lo tanto, una perspectiva situacionista de la misma.

En la actualidad, desde las perspectivas conductuales se utilizan técnicas de observación, autoinformes e instrumentos psicofisiológicos para realizar un análisis funcional de la conducta y sus variables ambientales y personales. Pero a medida que el panorama se amplía, se diversifica y se encauza más hacia la evaluación de variables internas, se hace más difícil hablar de método básico y de unas técnicas características de tal forma que en la actualidad, el modelo estaría caracterizado por su multimetodismo, a veces tan sorprendente como el empleo de tests proyectivos por los psicólogos conductistas (Schorr, 1993).

Poseemos excelentes estudios sobre las características y aportaciones del modelo conductual. Entre ellos merecen destacarse los de Barrios y Hartman (1986), Staats (1986), Barrett, Johnston y Pennypacker (1986), Nelson y Hayes (1981, 1986). Y, en España, los de Silva (1983, 1985, 1989, 1993ab), Mayor y Labrador (1984), Moreno (2003), Fernández-Ballesteros (1979, 1983a, 1992a, 1994, 2004a), Pelechano (1988), Martorell (1988), Tous (1989a), y Echeburua (1993a) entre otros. No obstante, cabe subrayar el excelente análisis de la evolución del enfoque conductual realizado por Forns et al. (1990, 1991) y Kirchner et al

(1998), que ha representado una guía para la selección de los datos más significativos presentados en este apartado. En los 90, se abrió un período en el que se llevaron a cabo diversos análisis y críticas respecto al estado de la evaluación conductual que permitieron nuevas configuraciones más integradoras en el seno de esta perspectiva, que ponen el énfasis en no perder su precisión metodológica.

No obstante, es importante señalar que la evaluación conductual aporta a todo el proceso diagnóstico un mayor rigor científico: Elige el método hipotético-deductivo, realiza inferencias de bajo grado, pretende mayor precisión en la medida y establece un mayor control tanto en la aplicación de los programas de tratamiento como en el resultado final, haciendo hincapié en que sus hallazgos o conclusiones puedan ser verificables. Fernández-Ballesteros (1994b, 2004a) realizando una reflexión sobre las características de la evaluación conductual, destaca que el objetivo de la misma es el cambio, se enfatiza la determinación ambiental del comportamiento, aunque se acepta la multicausalidad de la conducta anormal, la evaluación se plantea desde una perspectiva idiográfica, el triple sistema de respuesta como base de la definición operativa del problema de conducta, y la utilización de múltiples fuentes de datos y métodos en la recogida de información.

Existe en la actualidad un reto en la evaluación en salud mental debido al solapamiento encontrado entre síntomas psiquiátricos y conductuales, por lo que para la evaluación conductual se plantea mayor rigor en el proceso de diagnóstico. Para ello se ve la necesidad de mejorar el entrenamiento de los evaluadores, la utilización de instrumentos de evaluación más rigurosos y actualizados, así como la inclusión de la comorbilidad de síntomas para definir los términos del problema en su proceso de evaluación conductual (Holden y Gittlewen, 2007; 2009). Este solapamiento entre síntomas psiquiátricos y conductuales así como las dificultades de deslindar la comorbilidad son especialmente importantes en sujetos con retardo mental por lo que en este ámbito se debería ser especialmente cuidadoso (Hemmings, Gravestock, Pickard y Bouras, 2006; Rojahn, Matson, Naglieri y Maeville, 2004).

Una excelente revisión actualizada sobre la evaluación conductual en niños y adolescentes ha sido realizada por Reitman, Christiansen y Synder (2008). El interés por las intervenciones educativas derivadas de dicha evaluación se han expuesto por varios estudiosos del tema (Gresham; 2004; Kazdin, 2005; Kelly, 2003; Kwack, Ervin, Anderson y Austin, 2004; Mash y Hunsley, 2005; Orvasches, 2006).

Diversas voces se levantan a favor de una evaluación conductual más integrativa y menos analítica, que proponga un diagnóstico y medidas de intervención asociadas al mismo. Además, se hace hincapié en el interés de formularse en términos más comprensivos-integrativos lo que facilitaría entender y asumir dicha evaluación-intervención por parte de profesionales clínicos, trabajadores sociales, educadores y padres (Cipani y Schok, 2007; Craig y Beverly, 2008).

3.5.1. ENFOQUE CONDUCTUAL-RADICAL: CONDUCTISMO DE PRIMERA GENERACIÓN

Orígenes del modelo

Los orígenes de la evaluación conductual están estrechamente vinculados a las propuestas del conductismo de la primera generación, es decir, al conductismo clásico y al conductismo operante. Los trabajos de Watson o Skinner son representativos de esta etapa.

Presupuestos teórico-conceptuales básicos

En este enfoque se acepta únicamente como núcleo de análisis la conducta directamente observable, explicable por fenómenos de asociación entre los estímulos recibidos y las respuestas producidas, y entendidas bajo el concepto de relaciones funcionales de matiz adaptativo establecidas con la situación. Para Skinner la conducta sólo puede ser un efecto producido por las variables ambientales: Una persona no actúa sobre el mundo, el mundo actúa sobre ella (Skinner, 1971). Esta evaluación conductual-radical otorga un prioritario papel al ambiente externo como elicitor de la conducta, rechazando el estudio de toda variable de signo intrapsíquico. Si bien no niegan la existencia de posibles variables de persona como determinantes de la conducta, sin embargo discuten la adecuación de estas variables como objeto de estudio.

Objetivos de evaluación y unidades de análisis

Este modelo de evaluación se plantea por objetivo el control de la conducta, así como su predicción. La medida o el análisis de la conducta está estrechamente vinculado al tratamiento, ya que la evaluación conductual propiamente dicha consta de dos fases: (1) la de pretratamiento, en que se establece la línea base y las relaciones funcionales de la conducta con las variables que la controlan o mantienen; y (2) la de postratamiento en la que se evalúa el cambio por efecto de la intervención operada.

La conducta es entendida como muestra representativa del repertorio conductual que cada sujeto posee, y debe ser captada y explicada sin recurrir a constructos hipotéticos de los que pudiera ser el signo. Todas las variables intrapsíquicas o todo aquello proveniente de organismo que no pueda ser expresado en actos o movimientos objetivables y medibles, debe ser olvidado, “queda fuera de estudio y por lo tanto fuera del alcance del Psicodiagnóstico-Evaluación Psicológica” (Pelechano, 1988a). El objeto de análisis es la conducta del sujeto, ideográficamente considerada, y supuestamente elicitada por las situaciones. La precisión en la medida, tanto de la conducta manifiesta como de las situaciones que la elicitán, requiere que éstas sean definidas operacionalmente, a fin de que la identificación de la relación funcional que se establezca entre ambas tenga valor objetivo (Forns et al., 1990, 1991; Kirchner et al., 1998).

Los objetivos prioritarios de este modelo son la predicción y el control de la conducta y su predicción. Además desde esta opción, el análisis conductual está muy ligado al tratamiento. La evaluación conductual consta de las siguientes fases: a) la de pretratamiento o evaluación, en esta fase se establece la línea base y las relaciones funcionales de la conducta con las variables que la controlan o mantienen; y b) la de postratamiento en la que se evalúa la efectividad de la intervención (Kirchner et al., 1998).

Desde esta perspectiva se sostiene que la consistencia conductual no se deriva de las características del sujeto, sino de las propiedades de la situación. Por ello, el tratamiento se dirigirá a modificar las variables ambientales para establecer nuevas relaciones funcionales entre el estímulo y la respuesta.

Metodología y técnicas

Desde este enfoque se asume el método experimental para conseguir el máximo control y someter cualquier evento a las normas de objetividad empírica, con lo que se rechazan los métodos correlacionales. Los dos métodos fundamentales por los que se aboga son la observación y la experimentación.

La observación se utiliza para obtener y cuantificar datos, que sirvan para identificar los elementos o condiciones que eliciten o mantienen una conducta. A fin de operativizar el método observacional se han ideado numerosas técnicas de recogida de datos, entre las que destacan: los registros narrativos, las escalas de apreciación, los códigos de conducta, los autorregistros, las escalas, inventarios o cuestionarios que pueden ser hetero-descriptivos cuando lo que interesa evaluar es la conducta de un niño. Los instrumentos de observación o técnicas de análisis deben poseer las condiciones de validez y fiabilidad y, además, deben tener sensibilidad para captar la ocurrencia y frecuencia de aparición de la conducta, deben poder detectar la ocurrencia de la misma en más de un contexto y, finalmente, deben detectar la relación existente entre la conducta estudiada y otras conductas (precisión convergente). La metodología de análisis requiere de la identificación de la conducta problema, a la vez que, la de sus antecedentes y consecuentes, así como la determinación de sus características en cuanto a frecuencia, intensidad, duración... en relación a personas o contextos (Kirchner et al., 1998).

La experimentación se centra en la comprobación de los efectos que produce en la conducta del sujeto la manipulación de las variables situacionales-ambientales identificadas. El cambio de conducta puede lograrse mediante la dispensa de refuerzos positivos, de refuerzos negativos o el no reforzamiento de la conducta. Mientras la observación se utiliza en la fase de evaluación, la experimentación con el control y manipulación de variables se realiza, preferentemente, en la fase de modificación o terapia.

Según recogen Kirchner et al. (1998), el proceso de evaluación/modificación de conducta tiene varios momentos relevantes: 1) identificación de las variables que controlan la conducta alterada o variables independientes; 2) identificación y delimitación de la conducta problema (intensidad, duración, frecuencia...); 3) identificación de las consecuencias de las conductas refuerzos que recibe la conducta problema y tipo de refuerzos sustitutorios, 4) evaluación del cambio operado tras la manipulación de la variable independiente; 5) evaluación del mantenimiento del cambio; y 6) evaluación de la aparición de conductas sustitutorias.

Aspectos metodológicos y técnicas empleadas

Se rechazan los métodos correlacionales planteándose como bases metodológicas la observación y la experimentación. La observación tiene como característica el rigor y la meticulosidad en la cuantificación de los datos, como primer paso hacia la identificación de los elementos o condiciones que eliciten o mantienen la conducta. La experimentación se centra en la comprobación mediante manipulación de las variables independientes (ambientales-situacionales) identificadas, de los efectos que tales cambios producen en la variable dependiente o conducta del sujeto.

Los instrumentos de observación o técnicas de análisis que utilizan, además de poseer las clásicas condiciones psicométricas de validez y fiabilidad, deben ser precisos y sensibles para detectar: 1) La ocurrencia de un fenómeno o conducta; 2) La temporalidad de la ocurrencia o la

repetibilidad de un fenómeno; 3) La ocurrencia de un fenómeno bajo distintas condiciones estimulares, es decir, en más de un contexto; 4) El parámetro que mejor defina la conducta sometida a análisis; y 5) La relación existente entre el fenómeno estudiado y otras conductas (Forns et al., 1990, 1991; Kitchner et al., 1998).

Esta metodología requiere la identificación y definición operacional de la conducta problema, así como su análisis funcional, es decir, el análisis de los estímulos antecedentes y consecuentes; y su análisis topográfico, o sea, intensidad, frecuencia, duración....

Ámbitos de aplicación

Inicialmente comenzó a aplicarse en el ámbito del laboratorio, pero posteriormente se extendió a situaciones clínicas (tratamientos de disfunciones sexuales, fobias), educativas (prerrequisitos conductuales para establecer determinados aprendizajes, técnicas de adquisición y modificación de hábitos de conducta...), deportivas, etc.

3.5.2. ENFOQUE MEDIACIONAL O DE IMPACTO MEDIATIZADO CONDUCTISMO DE SEGUNDA GENERACIÓN

Orígenes del modelo

Este modelo sigue considerando que las variables ambientales tienen un valor importante en la determinación de la conducta, pero incorpora variables intervinientes entre los estímulos y las respuestas del sujeto susceptibles de análisis y estudio. Según Forns (1990), sin abandonar la idea de la asociación E-R, se acepta la existencia de procesos centrales mediadores (preceptos, imágenes, ideas,...), que otorgan una mayor complejidad a la conducta y permiten explicar aprendizajes más elaborados.

El paso del modelo conductual radical al mediacional se atribuye a autores como Hull (1884-1952), Tolman (1886-1959) y Guthrie (1886-1959). Hull consideró que cuando se había logrado un objetivo una vez, este logro podía actuar como “estímulo anticipatorio” de éxito en una situación posterior parecida, lo que da lugar al establecimiento de cadenas asociativas determinantes de la respuesta actual, y esto llevaría a la aceptación de la internalización de cadenas de estímulos. Tolman introdujo la variable “propósito” como la variable interviniente de mayor relevancia, este autor planteó que la base de la conducta incluía los principios de motivación, juego, experiencias estéticas y necesidades de orden biológico. Guthrie, que redujo el estudio del comportamiento a la observación y predicción de movimientos, desarrolló principios del aprendizaje como el de la “contigüidad entre E-R” o el del “efecto de la práctica” (Kitchner et al., 1998; Forns, 1990)

Presupuestos teórico-conceptuales básicos

El conductismo mediacional se interesa por los procesos mediacionales (perceptos, imágenes, ideas...), es decir, por los procesos mentales que pueden ser descritos como estímulos encubiertos, respuestas encubiertas o consecuencias encubiertas que confieren gran complejidad a la conducta.

Dentro de esta posición mediacional pueden encontrarse nombres como Wolpe, Eysenck, Bandura y Rotter los cuales desde campos teóricos y empíricos diferentes explican la conducta

en base a variables intermediarias o constructos hipotéticos, capaces de ser sometidos a modificación mediante las técnicas adecuadas. Wolpe (1978/1981) a partir del presupuesto de la inhibición recíproca (generación de una respuesta antagónica como contracondicionamiento a la ansiedad), utilizó la desensibilización sistemática como técnica terapéutica en el tratamiento de la ansiedad. Junto con Eysenck consideró a la ansiedad como la principal variable mediacional. Además, Eysenck (1967) pone de relieve determinaciones de referencia biologicista. En su opinión, la explicación de la conducta debe ser tributaria de diversos factores (cognoscitivo, conativo, temperamental, y constitucional) cuya interacción funcional da lugar a diversos patrones de conducta. Y tal como consideró los constructos básicos de la personalidad (extraversión, neuroticismo y psicoticismo) estos están relacionados con variables del sistema nervioso central y están constitucionalmente determinados.

Otra importante contribución del paradigma mediacional se conoce con el nombre de Aprendizaje Social, cuyo máximo representante es Bandura (1962, 1965, 1971, 1977) y Bandura y Walters (1963). Bandura propone la alternativa del aprendizaje por observación de modelos o modelado, destacando el valor social del aprendizaje por imitación como variable interviniente o mediacional. Toda teoría que acepte el aprendizaje por observación, modelamiento o imitación, requiere la aceptación de la representación interna de las acciones que puede ser disociada de su producción actual (Annett, 1989). Entre los años 60 y 70 van a surgir una serie de autores considerados como los máximos representantes del Aprendizaje Social (Bandura, 1962, 1965, 1971; Bandura y Walters, 1963; Rotter, 1954). Bandura y Rotter coinciden con Skinner en que la conducta se aprende y en que el reforzamiento es decisivo para el aprendizaje, pero difieren de él en su interpretación de la naturaleza del reforzamiento (Schultz y Schultz, 2003).

La teoría del aprendizaje social introduce la variable cognitiva como mediatizadora del impacto directo del ambiente sobre la conducta, ya que en opinión de Bandura (1981), la mayor parte de las influencias externas afectan a la conducta a través de procesos cognitivos intermediarios. Olivar (2001) señala que el concepto de autoeficacia de Bandura (1999) tiene una gran aplicación en el ámbito escolar, ya que la autoeficacia percibida se refiere a las creencias en las propias capacidades para organizar y ejecutar los cursos de acción requeridos para manejar situaciones futuras. Las creencias de los niños en relación a su eficacia se desarrollan a través de experiencias de dominio, de las experiencias vicarias presentadas por los modelos sociales, de la persuasión social, y de los estados psicológicos y emocionales al juzgar sus capacidades. En 1987, Bandura cambia el nombre a su teoría del "aprendizaje social" por el de teoría "cognitivo social", al concebir el aprendizaje como adquisición de conocimientos mediante el procesamiento cognitivo de la información. Según la teoría cognitivo social de Bandura (1999), los niños deben desarrollar destrezas para regular los determinantes motivacionales, afectivos, sociales y cognitivos de su funcionamiento intelectual.

En esta línea, Rotter (1954) siguiendo un enfoque situacionista, piensa que el refuerzo es importante en la determinación de la conducta, pero pone el énfasis en la representación que del mismo efectúa el sujeto en base a su experiencia pasada. La teoría de la personalidad de Rotter descansa sobre cuatro constructos: Potencial de conducta (ocurrencia de la conducta en una situación determinada en relación a un refuerzo), Expectancia (representación de la probabilidad de obtener cierto refuerzo), Valor del Refuerzo (para el sujeto), y Situación Psicológica. A juicio de Tous (1986) estos cuatro constructos pueden ser considerados variables intermediarias cognitivas, dado que operan con información proposicional (referencial

o virtual) en un nivel de abstracción bajo o medio. Para Rotter, la principal variable mediacional es la expectancia de control, y en relación con el constructo de expectancia desarrolló el concepto de "locus de control", diferenciando el locus de control interno (atribución a causas personales) del locus de control externo (atribución de los acontecimientos a fuerzas externas).

Como señalan McFall y McDonel (1986), el punto de mayor fricción entre mediacionales y situacionistas se centraría en la consideración, por parte de los primeros, de las influencias acumuladas del ambiente en términos de cambios orgánicos que mediatizan el efecto de los estímulos.

Objetivos de evaluación y unidades de análisis

El conductismo mediacional o de segunda generación se interesa por los procesos intermedios entre los estímulos y las respuestas por lo que su formulación teórica es E---(R-E)--R (Kirchner et al., 1998). El objetivo principal de la evaluación consiste en obtener datos específicos de la conducta del sujeto en los que basar la intervención. Se amplían las variables de estudio al tenerse en cuenta los estímulos elicitadores que mediatizan las respuestas del sujeto (Moreno, 1997).

El objetivo primordial de la evaluación desde estos enfoques mediacionales es la de predecir, instaurar, modificar o cambiar la conducta, es decir, controlarla. El enfoque mediacional amplía las unidades de análisis con respecto al conductismo radical. Hull plantea que deben ser controladas las siguientes variables intervinientes: la fuerza de los hábitos, la necesidad, el dinamismo o intensidad del estímulo y el incentivo motivacional. Rotter evalúa el "locus de control", y Bandura elige como unidades básicas de análisis el efecto que producen las variables independientes (influencias externas o eventos) en las variables cognitivas mediacionales, y cómo éstas repercuten finalmente en la conducta posterior. Como señala Kirchner (1989), las unidades de análisis del modelo mediacional han sido ampliadas respecto a los modelos más radicales. La ampliación va referida a la inclusión de los fenómenos encubiertos y las matizaciones que afectan tanto a las variables estímulos como a los consecuentes de la conducta (Forns et al., 1990, 1991).

Aspectos metodológicos y técnicas empleadas

En la misma dirección que los conductistas radicales, desde el enfoque mediacional se plantea la necesidad de rigor en el análisis y en la interpretación de la conducta humana. La ampliación de las variables de análisis, variables cognitivo-emocionales, trae consigo la aceptación del uso de los autoinformes verbales como medio de acceder al conocimiento de las expectancias y al valor del refuerzo. Se produce una gran elaboración de cuestionarios referidos a temáticas muy diversas que asumirán la idea de la valoración referida a criterio como elemento de máxima validez. Los presupuestos del aprendizaje social, especialmente el modelado (en vivo o vicario), se utilizan preferentemente como técnicas de intervención en la reducción de respuestas fóbicas y ansiosas con resultados alentadores. También la propuesta de Wolpe y Eysenck de la terapia basada en la inhibición recíproca de la ansiedad con cualquier otra respuesta incompatible, o la desensibilización sistemática como una de las técnicas más utilizadas en modificación de conductas ansiosas (Kirchner, 1989).

El método básico empleado por los investigadores de este modelo es el experimental con el objetivo explícito de predecir, instaurar y controlar (modificar) la conducta a partir de la

manipulación de las variables independientes (Kirchner et al., 1998). La ampliación de las variables de análisis a aspectos cognitivo-emocionales da lugar al uso de las técnicas de autoinformes, como medio de conocer las estas variables mediacionales. Se produce una gran elaboración de cuestionarios referidos a temáticas muy diversas. También se crean técnicas de tratamiento específicas aplicables a problemas concretos como la desensibilización sistemática y la exposición en el caso de la ansiedad, el modelamiento y la imitación en la instauración de conductas sociales, etc.

Ámbitos de aplicación

Su campo de aplicación es eminentemente clínico, aunque tiene aplicabilidad con objetivos preventivos. La técnica de desensibilización sistemática, o las técnicas de relajación se utilizan para la reducción de la ansiedad. Las técnicas de modelamiento y de imitación se aplican especialmente en contextos educativos y ámbitos sociales, así como en campañas comunitarias (prevención de accidentes...).

3.5.3. ENFOQUES INTEGRADORES E INTERACTIVOS: CONDUCTISMO DE TERCERA GENERACIÓN

Para algunos autores las unidades de análisis deben ser los atributos o rasgos intrapsíquicos de las personas, para otros los atributos de la situación, es decir, el ambiente, mientras que otros admiten que ambas variables determinan la conducta. Esta es la postura interactiva desde la que se asume que “la conducta es función de un proceso continuo de interacción multidireccional o feedback entre el individuo, la situación y sus encuentros (Magnusson y Endler, 1977). Este enfoque enfatiza “que la persona y la situación no son entidades separables, sino que forman un todo indivisible, un sistema continuo (Magnusson y Allen, 1983).

Estas teorías consideran como contribuciones la interacción aptitud x tratamiento, a partir de modelos dialécticos, transaccionales y del desarrollo. Son teorías dinámicas funcionalistas que tienden a la elaboración de un modelo multidimensional no mecanicista para explicar la interrelación persona/medio como interacción. Desde este enfoque se presenta la teoría interaccionista como una teoría integradora de algunas de las características de las restantes teorías de la psicología de la personalidad, pero su análisis de las restantes teorías es por una parte simple y por otra no constata que estas se hayan agotado en su capacidad de explicar o simplemente describir los fenómenos.

Desde la perspectiva interaccional, sujeto y situación se determinan mutuamente. El sujeto interactúa en el contexto al que otorga significación en función de su historia y características propias, y el contexto, a su vez, influye y determina la conducta del sujeto. Los modelos incluidos dentro de esta posición interactiva, en opinión de Kirchner et al. (1998), comparten los siguientes postulados:

- 1) La conducta se entiende como un proceso dinámico de interacciones multidireccionales.
- 2) El sujeto tiene un rol activo, es agente de sus acciones y actúa en base a su historia de aprendizaje. Y, a la vez, el contexto es también activo y provoca modificaciones en la conducta del sujeto.
- 3) Los factores cognitivos son básicos en la determinación de la conducta.
- 4) La situación percibida y la real tienen el mismo peso en la determinación conductual.

5) La conducta es situacionalmente específica y cross-situacionalmente general.

Las principales dificultades del modelo interaccional están en la explicación del cómo se produce la interacción y en el control de las múltiples variables que pueden estar influyendo y/o modulando cada uno de los elementos en interacción. A pesar de ello, el modelo es rico conceptualmente y supera la dicotomía de las posturas personalísticas y situacionistas. No obstante, algunos autores como Staats (1986) consideran que el interaccionismo combina más que unifica los modelos previos de explicación de la conducta, siendo necesario explicar cómo se estructuran nuevamente personalidad y aprendizaje a lo largo de la cadena evolutiva.

Una aportación de esta perspectiva tiene que ver con el debate consistencia de la conducta (personalistas) - especificidad de la conducta (conductistas radicales), y actualmente se reconoce que la consistencia absoluta a través del tiempo y de las situaciones no es posible. Otra aportación hace referencia a los aspectos metodológicos del modelo. Las dificultades técnicas que suponen la captación del complejo bucle interaccional de persona, conducta y situación, ha potenciado nuevas vías de análisis que se concretan en la elaboración de diseños de explicación multivariantes y de estructura causal (Log Linear Analysis and Lag Sequential Analysis). A continuación se tratan de forma específica los modelos conductual-cognitivo y conductismo paradigmático, por considerarse como los más representativos.

3.5.3.1. MODELO CONDUCTUAL-COGNITIVO

Orígenes del modelo

Los enfoques mediacionales abrieron el conductismo radical a procesos intermedios entre el estímulo y la respuesta, dando lugar a que en su propio contexto se fuera generando un movimiento denominado conductual-cognitivo cuyo foco de interés era el estudio de los procesos encubiertos, especialmente de orden cognitivo, a partir del paradigma experimental. Es justamente dentro de este marco conceptual donde la evaluación conductual se consolida.

La inclusión de lo cognitivo como variable de evaluación es antigua, pero dentro de la tradición concreta de la evaluación conductual pueden destacarse algunos pioneros, como Lang (1968) quien introduce las respuestas cognitivas (entendidas como todo aquello que piensa o siente el sujeto), como unidad de análisis en el diseño conductual. Pese a ello hay que señalar que en la práctica lo cognitivo quedaba reducido a lo expresado verbalmente por el sujeto. La progresiva dimensionalización conferida a lo cognitivo dará lugar a que dentro del conductismo se opere lo que algunos han denominado como una revolución cognitiva (Mahoney, 1977).

Autores representantes de esta concepción teórica son entre otros, Ellis (1962), Cautela (1968), Mahoney (1974, 1977), Beck (1976), Meichenbaum (1976). También podría considerarse a Bandura (1977) que si bien es considerado por algunos como un autor más vinculado al paradigma mediacional, otros analizando la evolución que siguió la obra de Bandura lo ubican en el modelo conductual-cognitivo.

Beck (1976) es el autor de la terapia cognitiva cuyo objetivo se centra en analizar los patrones de pensamientos erróneos que puedan dar lugar a conductas poco adaptativas. La terapia cognitiva se focaliza en el aquí y ahora, centrándose en las experiencias internas (mentales) del sujeto, tales como pensamientos, deseos, aspiraciones, etc. Para Beck es clave la idea de que las acciones de las personas sólo representan hechos que interpretamos. En la

obra de Beck la principal variable de análisis es el pensamiento, controlando el pensamiento se controla la conducta.

Ellis (1962) es el autor de la terapia racional-emotiva consistente en identificar los pensamientos irracionales que mantienen la conducta desajustada, y en reestructurar ese tipo de pensamientos reemplazándolos por principios lógicos. Las creencias racionales se traducen en consecuencias emocionales y/o conductuales adecuadas, mientras que las irracionales producen sentimientos o conductas inadecuados. La teoría de Ellis, conocida como esquema A-B-C, mantiene que ante un suceso real (A) se generan una serie de pensamientos, creencias o autoverbalizaciones (B) que conducen a la expresión de emociones y conductas que son consecuentes (C) de A y B.

Meichenbaum (1977) desarrolló la técnica del entrenamiento autoinstruccional consistente en enseñar a los sujetos a instruirse adecuadamente a sí mismos sobre cómo enfrentarse a situaciones problemáticas. Considera que los sujetos pueden hablarse a sí mismos y conseguir conductas adaptadas como consecuencia de la fuerza instruccional de la conducta verbal que actúa como modelado cognitivo.

Presupuestos teórico-conceptuales básicos

El núcleo conceptual de ese posicionamiento se centra en la importancia dada a los procesos cognitivos o simbólicos en el desarrollo, mantenimiento y modificación de la conducta, admitiéndose la existencia de una interacción recíproca entre las variables del ambiente, las del organismo y las de la propia conducta. El ambiente quedará relativizado o cognitivizado, adoptando una doble dimensionalidad: una como ambiente real y otra como ambiente percibido. Mayor y Labrador (1984) sintetizan en tres puntos las principales concepciones de esta perspectiva:

- 1) Las personas desarrollan conductas (adaptativas y desadaptativas) y patrones afectivos a través de los procesos cognitivos.
- 2) Se supone isomorfismo entre los procedimientos funcionales que activan a los procesos cognitivos y los del aprendizaje humano en el laboratorio.
- 3) La tarea de terapia-evaluación se centra en el análisis de los procesos cognitivos desadaptados y en la facilitación de experiencias de aprendizaje que puedan alterar las cogniciones y los patrones de conducta a ellos vinculados.

Objetivos de evaluación y unidades de análisis

Los objetivos fundamentales de la evaluación conductual-cognitiva se centran en identificar los patrones de pensamiento irracional o incorrecta autovaloración (Beck); la peculiar interpretación del medio que efectúa un determinado sujeto (Ellis); la escasez de expectativas, tipo de autoverbalizaciones y autoinstrucciones (Meichenbaum); o el tipo de habilidades y estrategias que posee para enfrentarse con eficacia a situaciones problemáticas (Mahoney, Spivack...). En síntesis, se trata de detectar y analizar qué variable de tipo interno-subjetivo está elicitando o controlando la conducta del sujeto.

Se consideran como unidades de análisis las variables de respuesta, las variables del ambiente y las variables del organismo. Con respecto a las variables conductuales se atenderá tanto a las conductas motoras, como también a las cognitivas y las fisiológicas. Es el análisis de

estas variables internas de tipo cognitivo (pensamientos, motivaciones, expectativas...) lo que diferencia fundamentalmente, a mi entender, la evaluación conductual-cognitiva de la conductual más radical y mediacional. A estas variables cognitivas en el diseño experimental se les puede otorgar valor tanto de variables independientes (etiológico), dependientes (conducta-problema), como mediacionales.

Así mismo, siguiendo la tradición del análisis funcional, se contemplarán los consecuentes de dicha conducta, ya que pueden ser los factores que la están manteniendo o controlando. Por lo que se refiere a las variables del ambiente, se atenderá tanto a los estímulos del ambiente real como percibido. Finalmente se tomarán como posibles unidades de análisis las variables procedentes del organismo, ya que pueden contribuir a la elicitación y mantenimiento de la conducta-problema (Nelson y Hayes, 1986). Estas variables hacen referencia tanto a aspectos biológicos como a la historia de los aprendizajes del sujeto (Márquez, 1987). Las competencias conductuales o cognitivas (Mischel, 1968/1973), los repertorios básicos de conducta (Staats, 1986), los sistemas autorreguladores (Meichenbaum, 1976), las ideas irracionales (Ellis, 1962), serán diversos elementos cognitivos relacionados con la particular historia de aprendizajes del sujeto que pueden estar incidiendo en una determinada respuesta.

Kirchner et al. (1998) señalan los siguientes elementos como característicos de este modelo: a) Todas las terapias se basan en el presupuesto de que los procesos cognitivos juegan un papel crítico en el mantenimiento de la conducta anormal; b) Las verbalizaciones incluyen interpretaciones distorsionadas, irracionales o inadecuadas de los acontecimientos externos actuales; c) Se puede concienciar al sujeto de la asociación entre sus problemas y sus verbalizaciones o expresiones no adecuadas; y d) Si el sujeto aprende a examinar estas expresiones y ponerlas en duda, irá tomando cuerpo la validez y racionalidad de los acontecimientos externos, y como consecuencia pensamientos más adecuados irán sustituyendo a los no adecuados, lográndose una conducta más adaptativa, precisa y racional.

Aspectos metodológicos y técnicas empleadas

El método fundamental de investigación inherente al modelo es el hipotético-deductivo, en su vertiente experimental y la observación directa de la conducta en su medio natural sigue siendo la técnica metodológica por excelencia. Pero además de la utilización de la observación, desde esta vertiente se han creado diversas técnicas evaluativas para analizar las construcciones internas. En este sentido se recupera la técnica de la introspección como medio de análisis de los contenidos no observables directamente operativizada mediante autoinformes y técnicas de autorregistro. Sin embargo, la utilización de estas técnicas no supone una vuelta al introspeccionismo al estar controlada objetivamente con métodos adecuados. No obstante, son evidentes las dificultades metodológicas inherentes al paradigma cognitivo, especialmente la objetivación de los procesos cognitivos que ocurren y, por ello, son accesibles solamente a través de inferencias a partir de los datos conductuales.

Así, en este enfoque se observa cierto enfrentamiento entre las exigencias científicas de objetividad y la presión del objeto de estudio (conductas subjetivas). Como señala Pelechano (1988) este problema fue resuelto por los modelos radicales del conductismo a base de adaptar el objeto al método. De hecho a nivel metodológico no se ha resuelto el problema de captar el bucle de interacciones entre las variables S-O-C (estímulo, organismo, conducta) pese a que se van depurando los procedimientos de análisis (Anguera 1990a, 1993, 1994b).

Clark (1988) distingue cuatro grupos de técnicas: 1) técnicas de registro (pensamientos en voz alta); 2) técnicas de producción (el sujeto anota sus pensamientos); 3) Técnicas de muestreo cognitivo (explicación de los pensamientos); y 4) técnicas de ratificación (cuestionarios estructurados), con los que no se intenta abarcar el estudio de toda la personalidad, sino parámetros concretos del comportamiento (depresión, pensamientos irracionales, ansiedad...).

Algunos autores (Barret, Johnston y Pennypacker, 1986) han considerado negativo la mezcla de técnicas. Frente al uso de técnicas más o menos propias del enfoque conductual (entrevistas estructuradas, registros psicofisiológicos, registros conductuales, etc.) ampliadas a la evaluación de conductas cognitivas (técnicas de pensamientos en voz alta, escalas de apreciación...), también es posible encontrar otras técnicas utilizadas por la evaluación tradicional - psicométrica (inventarios de personalidad, y otros tests psicométricos tradicionales). Sin embargo, para Fernández-Ballesteros (1994) el núcleo de la cuestión no estriba en la dispersión de técnicas, sino en cómo son usadas y en qué interpretaciones se hacen a partir de ellas.

Ámbitos de aplicación

Siguiendo la tradición del inicial análisis de conducta, su objetivo es la modificación de la conducta y, por lo tanto, su campo de aplicación prioritario fue el clínico. No obstante, y aunque en menor medida, también tiene aplicabilidad en otros ámbitos como el escolar, el industrial, o el de la evaluación de ambientes...

3.5.3.2. MODELO CONDUCTUAL PARADIGMÁTICO

Orígenes del modelo

El modelo de Staats, denominado conductismo paradigmático, aparece en la década de los sesenta como alternativa a la crisis del modelo conductual, como un intento de integrar posiciones contrapuestas. Staats (1963, 1975) elabora un marco teórico en el que integra los dos tipos de condicionamiento (clásico y operante), y aún en un único diseño variables situacionales, organísmicas, historia de aprendizajes y conducta manifiesta. Según Forns (1990), este diseño proporciona un marco de relación entre las teorías de la personalidad y las de la conducta.

El modelo propuesto por Staats y denominado Conductismo Paradigmático es otro de los que no acepta la linealidad en los determinantes de la conducta y que incluye en su formulación variables situacionales, variables de persona y variables de respuesta, además de variables de orden biológico. Se asienta sobre las asunciones derivadas del conductismo social (Staats, 1975; ver Fernández-Ballesteros, 1994e).

El conductismo paradigmático desde sus inicios (Staats, 1963, 1970, 1971, 1975) se implicó en la construcción de una teoría general del comportamiento humano que incluyera una teoría de la personalidad. Desde 1981 (Staats, 1981), momento en que presentó el Conductismo Paradigmático como una teoría unificada capaz de resolver las escisiones del propio Conductismo, ha ido perfilando su modelo, tanto desde el punto de vista conceptual como aplicado, teniendo importantes repercusiones en el área de la modificación de conducta y de la evaluación. El valor de su teoría, en sus propias palabras, está en que permite "tender un

punto de vista, el puente entre el conocimiento behaviorista y el conocimiento tradicional, y de esta manera proporciona un modelo metodológico para resolver el cisma entre el behaviorismo radical y la psicología tradicional” (Staats, 1989). Posteriormente, Tryon (1990) sugirió la red denominación del modelo a “Conductismo Psicológico”.

Presupuestos teórico-conceptuales básicos

Desde este modelo se establece que el comportamiento es función de las siguientes variables (Fernández-Ballesteros, 1994e):

- E1: Condiciones ambientales históricas, pasadas que forman parte de la historia de aprendizaje del sujeto.
- RBC: Repertorios Básicos de Conducta que han sido conformados a lo largo de la historia de aprendizaje del sujeto y que constituyen su personalidad.
- E2: Condiciones ambientales actuales que provocan, controlan o mantienen la conducta objeto de estudio.
- O1, O2, O3: Condiciones biológicas históricas que han contribuido en la constitución de los RBCs (O1), condiciones biológicas actuales que pueden afectar, actualmente, a los RBCs (O2), o condiciones biológicas actuales que interfieren en la recepción de las condiciones ambientales actuales (O3).
- C: Conducta objeto de estudio o problema de conducta por el que consulta el sujeto.

Una idea central de esta teoría paradigmática es la que hace referencia a los *Repertorios Básicos de Conducta* (RBC), que son complejas fórmulas conductuales aprendidas de forma acumulativa-jerárquica (de forma que las más sofisticadas apoyan a las más simples) y que determinan la conducta de los sujetos, la consistencia intrasujeto y las diferencias individuales. Fallos en este proceso acumulativo-jerárquico explicarían desde problemas de aprendizaje escolar hasta desajustes comportamentales. Estos *Repertorios Básicos de Conducta* son la esencia de la personalidad, este algo substantivo a lo que se refiere al hablar de la misma. Funcionalmente relacionados con condiciones ambientales del pasado pueden tener, en el presente, valor explicativo, otorgándoseles, por tanto, el status de variables independientes.

Desde que el individuo nace, y en base a los principios básicos del condicionamiento clásico e instrumental, va aprendiendo un sistema complejo de conductas en tres principales áreas que están interrelacionadas: la lingüística-cognitiva, la emocional-motivacional y la sensorio-motriz. Con los procesos madurativos, el sujeto va incrementando sus *Repertorios Básicos de Conducta* de modo que cada vez que experimenta y aprende una conducta nueva ésta no sólo está determinada por las condiciones estímulares de la situación actual, sino también por los repertorios previamente aprendidos (Forns et al., 1990, 1991; Kirchner, 1989).

Desde esta perspectiva se plantea que las circunstancias ambientales del pasado en las que se han ido adquiriendo los repertorios básicos de conducta, en conjunción con las circunstancias ambientales presentes, determinan la conducta actual del sujeto; ésta, a su vez, influirá en el desarrollo de repertorios básicos de conducta futuros y en posteriores condiciones ambientales que juntos incidirán en la conducta futura. Actualmente, Staats reconoce el valor de las variables orgánicas como mediadoras, confiriéndolas un valor fundamentalmente biologicista. Se trata de un modelo complejo, dentro de un marco interactivo, mediante el cual no sólo se puede predecir el comportamiento futuro de un sujeto en base al análisis minucioso

de sus *Repertorios Básicos de Conducta* (RBC), sino que también tiene aplicabilidad al campo de la terapia de conducta. La terapia va dirigida a cambiar los *Repertorios Básicos de Conducta*, no siendo necesario intervenir en la situación vital del sujeto.

Desde esta perspectiva, debido a que los *Repertorios Básicos de Conducta* son los elementos que configuran la personalidad de los sujetos, se defiende la consistencia conductual, explicando las diferencias individuales ante la misma situación en base a estos RBCs: "Los resultados indican que, cuando sujetos con diferentes características de personalidad, son expuestos ante la misma situación, responden de modo distinto. La situación (lo específico) no determina la conducta de los sujetos: lo hace su personalidad (sus sistemas emocionales-motivacionales)" (Staats, 1986).

Los presupuestos teórico-conceptuales que acabo de exponer inducen a reflexionar sobre la diferencia existente entre los RBCs (estables, consolidados en el individuo, determinantes de su personalidad, de su conducta, y de su consistencia conductual) y el concepto de rasgo de los modelos personológicos. Aunque estos últimos consideran que el rasgo tiene un fuerte componente innato-biológico y Staats habla de interacción con el ambiente en la conformación de los RBC sin darles a éstos ningún carácter de entidad intrapsíquica, sin embargo, no queda muy clara la diferencia, por ejemplo, entre un rasgo de introversión y una conducta introvertida en base al RBC del sujeto.

Fernández-Ballesteros y Staats (1992) consideran que el modelo se basa en la existencia de principios básicos de aprendizaje complementados por un principio acumulativo-jerárquico que establece las bases de aprendizajes complejos que ocurren históricamente. En su eje horizontal el modelo mantiene un análisis molar-diacrónico en el que se tienen en cuenta condiciones históricas, tanto biológicas como ambientales, así como las variables personales de sujeto que son entendidas como RBCs. En el eje vertical presenta un análisis funcional o sincrónico en el que se contemplan las relaciones funcionales entre condiciones ambientales actuales y la conducta objeto de estudio, así como las condiciones biológicas que inciden en la recepción de tales estímulos y en los repertorios básicos de conducta.

Objetivos de evaluación y unidades de análisis

La evaluación tiene por objetivo estudiar las variables situacionales y los Repertorios Básicos de Conducta, con la finalidad de estimular el desarrollo de habilidades (intelectuales, sociales...) promoviendo la modificación de la conducta del sujeto. Así mismo, la evaluación puede tener por objetivo una función de control de tratamientos de aprendizaje.

La evaluación psicológica se ocupa de la identificación de dos aspectos: (1) Análisis de las variables situacionales que inciden en la conducta; y (2) Evaluación del estado de los RBCs de un sujeto como determinantes de su comportamiento. Este segundo aspecto es el más específico y diferenciador del conductismo paradigmático, y se evalúan tres RBCs que están estrechamente relacionados entre sí:

- 1) El repertorio Lingüístico-Cognitivo: Se explica el desarrollo cognitivo de los sujetos íntimamente relacionado con las habilidades lingüísticas.
- 2) El repertorio Sensorio-Motor: Compendio de habilidades que permite al sujeto estructurar comportamientos tan variados como andar, escribir, cocinar..., siendo estas pautas motoras de gran importancia debido a que la mayor parte de los aprendizajes humanos descansan en el proceso imitativo o de modelamiento.

- 3) El repertorio Emocional-Motivacional: Se entiende que los estímulos que eliciten respuestas emocionales son considerados capaces de hacer que otros estímulos tengan valor emocional. O sea, si las palabras, por condicionamiento clásico, han sido asociadas a estímulos desencadenantes de respuestas emocionales, tales palabras pueden provocar directamente estas emociones.

Aspectos metodológicos y técnicas empleadas

El método utilizado desde este enfoque es el método naturalista-experimental debido a que estudia la conducta en condiciones naturales, introduce variaciones y controles para conocerla con precisión, y en diseños longitudinales. En este enfoque la técnica naturalista-experimental se aplica a la observación y análisis de la conducta de uno o varios sujetos a lo largo de extensos períodos de tiempo, en los que se comprueba cómo los aprendizajes previos van formando la base para otros aprendizajes más complejos, a través de un proceso acumulativo y jerárquico.

Respecto a las técnicas utilizadas, se emplean también tests psicométricos tradicionales, ya que, por ejemplo, los tests de inteligencia permiten conocer el estado de los aprendizajes y la coherencia del proceso acumulativo-jerárquico implícito en los *Repertorios Básicos de Conducta*. Así, se emplean este tipo de tests pero sus resultados son interpretados como muestras de lo que el sujeto es capaz de hacer a cierta edad. Para Staats y Fernández-Ballesteros (1987), estos tests recogen muestras de las habilidades conductuales básicas del niño en determinada edad, indicando que “en la teoría de la inteligencia del conductismo paradigmático se considera importante que tests como el Stanford-Binet midan muestras de habilidades conductuales básicas que la mitad de los niños a una determinada edad habrán aprendido. Y es debido a que las habilidades son esenciales para el futuro aprendizaje de los niños. En este sentido, estos autores observan que los tests de inteligencia son predictores del futuro éxito escolar del niño. Además, consideran que ante la pregunta de un test de inteligencia o personalidad, la respuesta que formula el sujeto ha sido asociada a la pregunta por condicionamiento clásico.

Para Staats y Fernández Ballesteros hay elementos de conocimiento derivados del campo tradicional en la aplicación de tests psicológicos que pueden ser incorporados al campo de la evaluación conductual. Por ejemplo, estos autores entienden que el informe verbal que emite un sujeto sobre sus emociones y pensamientos puede suministrar información adecuada sobre eventos internos (subjetivos) que no pueden observarse directamente.

Para esta integración del enfoque conductual y de los instrumentos tradicionales (cuestionarios de personalidad, de intereses vocacionales, de evaluación psicopatológica, tests de inteligencia...) resulta necesario un nuevo marco conceptual en el cual integrar esta unificación, y este enfoque, el del conductismo paradigmático, se plantea como modelo integrador que aporta el marco conceptual desde el que operar esta integración (Fernández-Ballesteros, 1994e).

3.6. EL MODELO COGNITIVO

El estudio de las variables cognitivas, como ponen de manifiesto Kirchner et al. (1998), está presente desde los inicios de la evaluación psicológica con el trabajo de Binet (1905), y, aunque posteriormente relegadas por la influencia del conductismo, siempre ha habido

estudiosos que las han tenido muy presentes (Piaget, Vygotsky, Chomsky, etc.). La psicología cognitiva en la actualidad incluye una variedad de modelos con diferencias, conceptuales, metodológicas y de objetivos, importantes entre sí. En un sentido general, el cognitivismo supone considerar las cogniciones y los procesos cognitivos como variables de análisis fundamental.

Orígenes del Modelo

Hacia los años 50 la concomitancia de varios hechos cristalizó en el surgimiento de un nuevo paradigma psicológico: el Cognitivismo o Psicología Cognitiva. La definición y concepto de la Psicología Cognitiva encierra en sí misma dificultades de consenso, pero en general se acepta que la Psicología Cognitiva es el enfoque cognitivo de la Psicología, es decir, el estudio de la conducta desde la perspectiva cognitiva. Así mismo ha sido definida como aquella subdisciplina de la Psicología relacionada con los procesos de alto nivel, tales como percepción, memoria, pensamiento, toma de decisiones, etc. (Lachman y Lachman, 1986).

La Psicología cognitiva ha recibido la influencia: (1) de los avances de la electrónica y con ella el impacto del ordenador volvió a legitimar el interés por los procesos y las estructuras internas porque ahora podían especificarse exactamente en programas de ordenador; (2) de las nuevas formulaciones de la psicolingüística planteadas por Chomsky, y (3) de la obra de Piaget.

Con frecuencia es considerado un paradigma más que un modelo, debido a la gran diversidad de planteamientos que se albergan en Psicología bajo la rúbrica de "cognitivo". En este sentido se integran enfoques tales como el procesamiento de la información (Estes, 1978), la psicolingüística (Chomsky, 1968), el neoconductismo (mediacional y cognitivista), el estudio de los procesos cognitivos desde la psicofisiología (Luria, 1979) la psicología genética (Piaget, 1959/1969), el papel integrador de la conciencia surgida de la interacción biológico-social (Luria, 1975, 1977, 1980), la teoría de los sistemas, o el enfoque de la orientación ecológica (Neisser, 1967, 1976). El intento de integración de aspectos procedentes de tan diversos planteamientos ha de pasar inevitablemente por un reduccionismo.

Presupuestos teórico-conceptuales básicos

El interés por los procesos cognitivos es inherente al desarrollo histórico del modelo, especialmente los procesos mentales superiores como la atención, percepción, memoria, lenguaje, pensamiento... Aunque es difícil integrar el consenso existente entre estos subparadigmas, se puede plantear que desde estas distintas perspectivas:

- 1) Consideran al sujeto como agente activo. El agente de la conducta no es un organismo vacío, ilimitadamente moldeable, sino que se define funcionalmente por una cierta organización, una determinada estructura o arquitectura. El sujeto de la psicología cognitiva es por naturaleza activo, es un sujeto que procesa la información, es decir, selecciona, codifica, transforma, recupera, etc. tanto la información proveniente del exterior como la que proviene de su medio interno.

- 2) Comparten el presupuesto de otorgar a las variables cognitivas (cogniciones, atribuciones, funciones mentales superiores como memoria, atención, percepción, lenguaje, pensamiento...), un valor de primordial importancia en la determinación de la conducta del individuo. Se considera que la conducta está determinada por una serie de procesos y estructuras mentales internas, es decir, es una función del mundo cognitivo de la persona. El comportamiento es explicado en base a una serie de estructuras internas mentales, dando especial importancia a los procesos mediadores o a estrategias cognitivas. Por ello se enfatizará el análisis de los procesos versus el análisis de los productos planteado desde otros modelos.
- 3) No se niega la importancia del aprendizaje, pero se admite cierta preprogramación y potencial biológico, así como una cierta independencia del funcionamiento cognitivo con respecto a las variables del ambiente. El sujeto cuenta con planes, propósitos, programas de acción que exceden a la simple conducta elicitada.

Como puede verse, el modelo cognitivo subraya, en la explicación del comportamiento, el estudio de las estructuras mentales (como la representación, los almacenes de memoria, las distintas fases en el procesamiento de la información), y más especialmente, los procesos o estrategias cognitivas que median entre los estímulos y las respuestas o ejecuciones (Fernández-Ballesteros, 1992a, 2004a).

Objetivos de evaluación y Unidades de análisis

El objeto de evaluación de la psicología cognitiva es el conocimiento de los procesos cognitivos mediacionales, en el sentido de cómo se organiza la situación estimular para responder propositivamente a la misma. En un sentido estricto hace referencia al conjunto de los procesos cognitivos, y en un sentido amplio a qué papel juegan los procesos cognitivos en la determinación de la conducta. El objetivo básico de una evaluación desde la perspectiva cognitiva es conocer como piensa el ser humano. Se trata de entenderlo como un sujeto que dispone de multiplicidad de planes y de esquemas y que puede usar de ellos en distintos momentos y de distinta forma; el objetivo es descubrir qué planes activa, cómo los usa, en qué casos y cómo los organiza y cómo mediante todo ello transforma la información (Forns, 1990).

Un objetivo común a todos estos subparadigmas es el de analizar los procesos cognitivos internos que permiten al sujeto procesar la información que le posibilitará la adaptación a la realidad. Si bien los procesos no son directamente observables pueden ser inferidos a partir del análisis del procesamiento de la información. Como indicaron Pelechano (1975) y Carpintero (1975), se pasó del estudio de "E" y "R" al estudio de la "situación estimular" y la "respuesta integrada y significativa". En definitiva, ha cobrado relevancia el papel activo del sujeto en la interpretación del estímulo y en la organización de la respuesta. En última instancia, desde el modelo cognitivo se pretende describir, predecir y explicar la conducta (Fernández-Ballesteros, 1992a; 2004a).

Desde esta perspectiva multiparadigmática, se plantea como unidad fundamental de análisis las estrategias usadas para resolver tareas cognitivas; analiza las correlaciones entre variables fisiológicas y operaciones cognitivas dando soporte a las diferencias individuales; se estudia la relación entre los diferentes estilos cognitivos (formas preferenciales de organizar e

interpretar la realidad) y diferencias individuales... Para evaluar estos nuevos objetivos, se están desarrollando y depurando algunas técnicas, aunque estas no son numerosas.

Aspectos metodológicos y técnicas empleadas

Se utiliza preferentemente la metodología experimental, pero no es el único método, ya que también se emplea la metodología observacional y correlacional. La interdisciplinariedad científica de este paradigma hace que se utilice el pluralismo metodológico que para Eysenck (1985) es la posición de más garantía. Los procedimientos o técnicas han derivado de la neuropsicología, la observación controlada e incluso la introspección. Los más usuales son la simulación y el análisis cronométrico.

La *simulación* supone una copia de la realidad, siendo su punto crítico la analogía entre las respuestas del ordenador y las del sujeto. Este procedimiento se ha aplicado a todo tipo de conductas. El *análisis cronométrico* es nuevamente activado por Sternberg en la década de los 60, introduciendo el método sustractivo con el método aditivo. Esta técnica se define como “la combinación del método que mide un proceso, y el método de los factores aditivos, que establece la existencia de etapas y de las variables correspondientes”. Estudios españoles llevados a cabo en la década de los 90 en torno a la evaluación fisiológica de la inteligencia han analizado las relaciones entre inteligencia psicométrica (medida con tests de inteligencia tradicional) y dos indicadores cronométricos: tiempo de inspección y latencia del componente P300 del registro de potenciales evocados cerebrales (Ferrando, Vigil, Lorenzo y Tous, 1993)

Las técnicas que se utilizan desde este enfoque son variadas, se emplean autoinformes (por ejemplo, pensamientos en voz alta), tests situacionales e instrumentos de evaluación fisiológica, que permiten apreciar la reacción de los sujetos ante determinadas tareas, baterías neuropsicológicas.... Además, se han elaborado algunas técnicas nuevas preferentemente de evaluación de la inteligencia (Forns, 1993), como:

- Mapas conceptuales para el análisis de la organización del conocimiento (Goldsmith, Johnson y Acton, 1991; Novak y Gowin, 1984/1988).
- Diversos cuestionarios contruidos en relación a contextos específicos de Sternberg (1985) y Sternberg et al. (1981) para la evaluación del conocimiento declarativo.
- La batería “*Learning Potential Assessment Device*” (LPAD) para evaluar el potencial de aprendizaje (Feuerstein et al., 1979).
- La batería de evaluación de la inteligencia, *Sternberg Triarchic Abilities Test* (STAT) (Sternberg, 1991) que está actualmente en fase de baremación.

Las diferencias en la conceptualización teórica y la utilización de los instrumentos de evaluación desde la perspectiva cognitiva, ha sido investigada recientemente por Eden (2008) concluyendo que aporta menos acuerdos (o más variabilidad) en la evaluación las diferencias en el constructo teórico que en el uso de herramientas diagnósticas.

Ámbitos de aplicación

Es difícil establecer nítidamente los límites del campo de aplicación, puesto que la confluencia de disciplinas científicas y de enfoques, dentro de una misma disciplina, ha hecho que su campo de aplicación sea amplísimo. Por citar los más cercanos se ha aplicado al campo

educativo-escolar, al campo de la psicolingüística, a las neurociencias, etc. En psicología cognitiva se abarca el estudio de los procesos cognitivos de memoria, aprendizaje, percepción... hasta su aplicación e intervención terapéutica (Maganto, 1995b).

El énfasis puesto por la Psicología cognitiva en que las diferencias individuales son debidas a que los sujetos procesan de modo distinto y parten de conocimientos distintos, es de gran importancia a la hora de mejorar la selección de objetivos de intervención y las predicciones que están en la base de la toma de decisiones en los contextos escolar y de selección de personal (Alonso Tapia, 1992).

3.6.1. LA EVALUACIÓN DESDE EL PROCESAMIENTO DE LA INFORMACIÓN

Orígenes del modelo

La fuente más cercana de esta nueva perspectiva proviene de la imbricación de la psicología experimental con la teoría y tecnología computacional, especialmente el subcampo de la Inteligencia artificial (Lachman y Lachman, 1986). Los avances tecnológicos en los campos de la cibernética, la informática etc. posibilitaron la creación de máquinas cada vez más complejas. Las primeras máquinas resolvían problemas de forma determinista, rutinaria, pero posteriormente se las dotó de mecanismos de retroalimentación que les permitía desarrollar respuestas a procesos de azar, y por último, se introdujeron los métodos computacionales capaces de modificar sus propios programas (Forns et al., 1990, 1991). La premisa básica de esta corriente consiste en el establecimiento de una analogía entre procesos computacionales y conceptos mentales.

Presupuestos teórico-conceptuales básicos

Tres grandes líneas de investigación han incidido en el desarrollo de este modelo (Forns, et al., 1990, 1991):

- 1) Los trabajos de Miller (1956) sobre la posibilidad de agrupar unidades de información y los de Broadbent (1970) sobre la capacidad de recepción y almacenamiento de la información, dieron pie a posteriores trabajos que permitieron la expresión del funcionamiento cognitivo mediante diagramas de flujo.
- 2) Bruner, Goodnow y Austin (1956) establecieron pautas regulares de resolución de problemas denominadas "estrategias" que explicaban cualitativamente el desempeño de una tarea por parte de los sujetos.
- 3) Newell y Simon (1972), creadores de lo que se ha denominado inteligencia artificial, establecen un paralelismo entre el ser humano y el ordenador considerando a ambos manipuladores de símbolos para propósitos generales. Así mismo, describen las características de un programa de resolución de problemas capaz de resolver problemas de ajedrez, acertijos, etc. Una derivación de esta última línea de investigación ha sido planteada recientemente por Neisser (1984) que rechaza la virtual igualdad de los canales de información de una máquina con los del ser humano, e insiste en la necesidad de romper con la simple concepción lineal del procesamiento, basado en el supuesto de que el sistema cognitivo humano posee una serie de estadios secuenciales para procesar la información, proponiendo sistemas de procesamiento simultáneo o en paralelo.

Este modelo es un enfoque conceptual que asume que el funcionamiento humano puede ser fácilmente conceptualizado y entendido en términos de cómo procesa y utiliza tanto la información del medio como la interna. El procesamiento de la información se refiere a un amplio paradigma conceptual y no a una teoría o modelo. El paradigma de procesamiento de la información aplicado al ser humano se define como un enfoque conceptual que presupone que el funcionamiento humano puede ser conceptualizado y comprendido productivamente en términos de cómo la información interna y la información externa son procesadas y utilizadas. A este enfoque le subyacen los siguientes presupuestos teórico-conceptuales:

- 1) Se considera al sujeto como un procesador activo de información procedente tanto del medio externo como del interno (estados somáticos y mentales) (Lachman y Lachman, 1986). La actividad humana será la resultante de la compleja interacción entre el contenido de una información, la elaboración que de ella efectúa el sujeto y los outputs que este retorna al ambiente (Ingram y Kendall, 1986).
- 2) En oposición al conductismo radical, las capacidades innatas y los procesos internos, como planes, estrategias, decisiones, etc., toman relieve (Lachman y Lachman, 1986).
- 3) La organización de la cognición descansa en la articulación dinámica de estructuras cognitivas (redes cognitivas, esquemas, marcos referenciales..., con sustrato neurológico y cierta estabilidad) y procesos cognitivos (componentes dinámicos del sistema que se refieren a las transformaciones y organización que experimentan las estructuras; hacen referencia al funcionamiento del organismo en un momento dado). Ambos constructos hipotéticos subyacen a todas las funciones cognitivas.

Esta perspectiva ha tenido importantes desarrollos dando pie a diversas investigaciones con la finalidad de estudiar los efectos de los niveles de procesamiento en pruebas de memoria (Menor y Sebastian, 1995), las relaciones entre inteligencia y velocidad o eficiencia del procesamiento de la información (Coscolluela, Andrés y Tous, 1992), o las relaciones entre personalidad, estrategias cognitivas y tiempo de inspección (Andrés, Vigil y Tous, 1989).

Objetivos de evaluación y unidades de análisis

Las unidades de análisis fundamentales de la evaluación psicológica en el modelo de P.I. son: (1) la identificación de las estructuras y procesos cognitivos, y (2) la determinación de las variables que influyen en unas y otros, y sus complejas interrelaciones. Debido a la relación entre cognición y conducta, la evaluación se centrará tanto en los elementos que componen la estructura cognitiva de un determinado sujeto y sus relaciones mutuas, como en el estudio de su particular proceso de transformación de la información, identificando sus estrategias de resolución. Este último aspecto, la evaluación de procesos, se constituye, a mi juicio, en uno de los puntos de mayor diferenciación con respecto a lo planteado desde otros modelos que se han centrado en la evaluación de productos.

Aspectos metodológicos y técnicas empleadas

En cuanto a los aspectos metodológicos este enfoque se caracteriza por la multiplicidad y combinación de técnicas que permiten tanto un acercamiento directo como indirecto a los

procesos objeto de estudio. Se adhieren al método experimental, pero lo combinan con métodos correlacionales y observacionales. Además, recuperan la introspección como método de investigación para acceder a lo cognitivo pero intentando mantener la objetividad necesaria que evite críticas por subjetivismo. Los datos de base siguen siendo objetivos, perfectamente medibles y controlables, y se concretan tanto en la situación estimular como en el patrón de respuestas. Se examinan las respuestas de los sujetos a las variables manipuladas y se efectúan inferencias lógicas y factibles sobre los procesos cognitivos subyacentes a los que se atribuye la responsabilidad de lo observado.

Otro dato metodológico de gran relevancia es el desarrollo y perfeccionamiento de métodos que hacen viable el análisis de datos cualitativos y que permiten conjugar la riqueza de una observación directa, con el rigor del número, paliando al propio tiempo el subjetivismo inherente al primero y el reduccionismo en que incurre el segundo (Anguera, 1978, 1981, 1983, 1985ab, 1986, 1988, 1989, 1990a, 1993, 1994b).

La tecnología empleada por este modelo es también muy variada. Emplean la observación y las técnicas de autoinforme (pensamientos en voz alta sobre los procesos mentales seguidos para resolver una tarea). Los análisis cronométricos, los programas de ordenador, los diagramas de flujo, técnicas de medida de tiempos de reacción, análisis de los movimientos oculares, mapeos cognitivos, procedimientos de medida de memoria, y técnicas taquiscópicas son otras técnicas que tienen utilidad en este modelo (Fernández-Ballesteros, 1992a; Forns et al., 1990, 1991; Kirchner, 1989; Mayor, 1980; Mayor, Suengas y González Márquez, 1993; Salvador, Pelegrina y Soler, 1988). No obstante, como matiza Tous (1989b) en el ámbito del tratamiento las técnicas adecuadas para llevar a cabo un tratamiento basado en el paradigma de procesamiento de la información, se diferencian de las técnicas propias de la modificación de conducta cognitiva en que no utilizan los autoinformes, o la evaluación directa de la información, socialmente significativa que dispone cada ser humano, sino técnicas indirectas generadas por el método experimental cognitivo.

Ámbitos de aplicación

Este enfoque está teniendo fuertes repercusiones principalmente en el área del aprendizaje escolar, en la clínica y en la investigación.

a) En evaluación clínica

Este enfoque intenta averiguar los procesos y mecanismos cognitivos que subyacen a determinadas alteraciones de conducta, la particular interpretación que efectúa el sujeto de su mundo y que lleva a una cognición desviada (alteración de los mecanismos selectivos de estímulos, lentificación del procesamiento de la información...). Así, se analizan las diferencias en el procesamiento de la información tanto en sujetos diagnosticados de disfunciones o alteraciones patológicas como en sujetos normales. Además, desde esta perspectiva intentan desarrollar medidas de riesgo o vulnerabilidad adecuadas como medio para detectar irregularidades en el procesamiento informativo. Una amplia revisión de estos aspectos ha sido planteada por Ingram y Hollon (1986) y Saccuzzo (1986).

Como ya había señalado previamente (ver perspectiva psicoanalítica), dentro del contexto clínico, y como una reciente línea de investigación, se están desarrollando diversos estudios que integran la psicología psicodinámica con la representación y procesamiento de información

emocionalmente significativa, desde la reformulación y comprobación experimental de conceptos psicoanalíticos. Algunos investigadores actuales de enfoque cognitivista están “focalizando su atención en el estudio de aspectos emocionalmente significativos de la conducta humana, aplicando modelos y técnicas cognitivas para medir y validar estos descubrimientos” (Strauman, 1994).

De hecho, la revista *Journal of Personality* en 1994 dedicó un monográfico cuyo objetivo fue reunir la opinión de varios científicos (Andersen y Baum, 1994; Crits-Christoph et al., 1994; Erdelyi, 1994b; Hughes et al., 1994; Strauman, 1994; Horowitz et al. 1994; Kihlstrom, 1994; Weinberger y Davidson, 1994; Wegner y Zanakos, 1994; Westen, 1994) que integrando ambos enfoques (psicoanálisis y psicología social-cognitiva) estudiaron cómo las personas representan y procesan información emocionalmente significativa, es decir, intentaron interpretar principios psicoanalíticos desde el marco teórico de la teoría del procesamiento de la información. Cada artículo incluido en este monográfico evidenció los nuevos hallazgos referentes a la naturaleza y vicisitudes de la información cargada afectivamente, presentándose variados diseños (experimental, correlacional, estudio de caso), métodos (cognitivos, psicofisiológicos, entrevista, autoinformes) y muestras (estudiantes, individuos en terapia).

Andersen y Baum (1994) reexaminan el concepto Freudiano de transferencia a la luz del modelo del procesamiento de la información, concluyendo que este fenómeno puede ser demostrado en el laboratorio a través del uso de técnicas cognitivas bien validadas. Desde su punto de vista, el objetivo de su investigación consiste en identificar y examinar empíricamente los procesos mentales básicos implicados en la transferencia (Andersen y Baum, 1994), definiendo la transferencia como un proceso que implica una representación de un otro significativo y una idiosincrática experiencia que puede aplicarse inapropiadamente con otros. En relación a este aspecto Kihlstrom (1994) señala que el estudio experimental de Andersen y Baum pone de manifiesto el fenómeno de la transferencia, sin embargo, estos resultados no apoyan la más radical hipótesis freudiana, porque este estudio no señala el rol de las relaciones arcaicas, ya que los otros significativos pueden no haber sido necesariamente los padres o figuras parentales. No obstante, la metodología experimental utilizada por Andersen y Baum muestra que las representaciones de gente significativa en la vida de uno realmente “tienen” las nuevas relaciones sociales. Kihlstrom subraya que siendo esto un gran avance, en el análisis final es importante conocer si esto es algo más que un caso de generalización de estímulos.

Crits-Christoph et al. (1994) estudian cómo los individuos tienden a percibir temas particulares en función de diferentes ambientes o escenarios. Presentan un nuevo método para evaluar temas significativos personalmente que ocurren durante la terapia, aportando evidencia estadística del concepto de repetición de temas interpersonales. Discuten la extensión del contenido de la repetición y su relación con el tipo de terapia. Horowitz et al. (1991) aplican un modelo integrativo de los procesos de control defensivos al estudio de una psicoterapia. Estudian los estados de la mente y su relación con procesos cognitivos de control. Los autores presentan evidencia de la existencia de relaciones entre estados emocionales que se caracterizan por específicos procesos defensivos. Utilizando técnicas de evaluación multidimensional subrayan procesos mentales en la psicoterapia.

Hughes et al. (1994) también presentan una nueva metodología para estudiar cómo la expresión de emociones y el uso de estrategias defensivas diferentes, podría ser reflejado a través de cambios momentáneos en la actividad del sistema nervioso autónomo. Los resultados de su estudio presentan sugerencias concernientes a la asociación entre procesos de control y

fisiología, con implicaciones para la salud psicológica y física. Weinberger y Davidson (1994) realizan diversos estudios empíricos sobre la represión, tratando de diferenciarla de otros estilos de inhibición de la expresión emocional. Usando experimentalmente tareas de laboratorio los autores concluyen la existencia de una diferencia entre ambos conceptos. Desde su punto de vista, los represores aparecen para inhibir su expresión de emociones con la finalidad de controlar su conciencia de afecto, más que en orden a maximizar la auto-presentación. En esta línea, Wegner y Zanakos (1994) intentan establecer que la inclinación hacia la supresión de pensamientos es un fenómeno distinto de otros regulatorios procesos cognitivos. Destacan la existencia de una relación entre la supresión crónica de pensamientos y variables relevantes clínicamente como obsesividad, reactividad emocional, y afectos depresivos y ansiosos.

Finalmente, Kihlston (1994) sustentándose en la psicología cognitiva y en la cognición social, presenta su propio punto de vista sobre la integración del psicoanálisis y la cognición social, sugiriendo que las dos perspectivas representan una rica fuente de inspiración e hipótesis. Para Kihlston (1994) el intento de vincular el psicoanálisis con la psicología científica es siempre interesante, pero existen muchas diferentes formas de psicoanálisis, y además muchos modos de hacer conexiones. Desde su punto de vista, la psicología cognitiva contemporánea ha redescubierto el inconsciente psicológico, pero la naturaleza de la vida mental inconsciente, es muy diferente de la que Freud planteó. Es demasiado pronto para decir si esta actual tendencia a la integración de enfoques acerca del sí mismo, del afecto y la psicopatología, finalmente liderará una más sofisticada y comprensiva comprensión de cómo la gente trata con aspectos emocionalmente significativos de sus vidas (Strauman, 1994).

b) En el contexto escolar

Desde este enfoque se intenta conocer qué estrategias y métodos utiliza el alumno experto en relación al novicio, indagar cada uno de los pasos que sigue el sujeto desde que recibe el input hasta que emite su respuesta, averiguar qué facilita y dificulta la propagación de la información...

Una de las áreas que ha generado mayor número de investigaciones es la relativa al análisis de la inteligencia, donde se operativizan diversas líneas de estudio: (1) estudio de correlatos cognitivos; (2) estudio de componentes cognitivos; (3) estudio de diversos aspectos del aprendizaje (entrenamiento cognitivo y potencial de aprendizaje); y (4) estudio de las estrategias y estilos de enfrentamiento y resolución de tareas. En este contexto, Sternberg (1985) estudia los distintos componentes cognitivos (identificación, comparación, etc.) necesarios para resolver un problema mediante la identificación experimental de los tiempos empleados en cada operación. Formula una teoría de los componentes cognitivos que ha sido integrada en la *Teoría Triarquica de la Inteligencia*. Una de las partes fundamentales pretende explicar las estructuras y mecanismos que subyacen a la conducta intelectual (subteoría componencial); otra, defiende la influencia socio-cultural y contextual de la conducta (subteoría contextual), y finalmente, otra postula el valor de la experiencia como matizadora de la habilidad cognitiva, y considera que el mejor método para captar tal habilidad consiste en la ejecución de tareas nuevas, en situaciones nuevas, mediante procesos de automatización (subteoría experiencial). Estas teorías de componentes y metacomponentes mentales de Sternberg, aunque discutidas, significan un paso adelante en este intento de dar una base teórica a la investigación psicométrica (Yela, 1995).

La aplicación de las teorías del P.I. a la evaluación psicológica ha dado lugar a otra línea de estudio denominada “entrenamiento cognitivo”. Desde este enfoque algunos investigadores (Campione, Brown y Ferrara, 1982, Campione y Brown, 1987, 1990) intentan identificar, mediante entrenamiento aquellas habilidades fundamentales para el aprendizaje y cuya ausencia o presencia determina los resultados de las ejecuciones. En diversos estudios analizan los componentes y operaciones cognitivas básicas que emplean los expertos versus inexpertos, con la finalidad de evidenciar las habilidades responsables del éxito o del fracaso en la ejecución de una tarea.

En similar dirección, cabe resaltar el trabajo de Feuerstein (1979, 1980, 1988) que diseñó un método de evaluación denominado *Learning Potential Assessment Device LPAD* con el que se estudia el cambio cognitivo que se produce en su sujeto como consecuencia de un entrenamiento breve. Su trabajo, además, ha modificado el concepto de evaluación psicológica (evaluación del cambio, evaluación para conocer el tipo de operaciones alteradas y determinar en qué fase se ubica esta alteración, la evaluación como una situación de enseñanza-aprendizaje en la que se valoran las respuestas semicorrectas...) y ha promovido la creación de nuevas técnicas de evaluación como el EPA (*Evaluación del Potencial de Aprendizaje*) de Fernández-Ballesteros, Calero, Campllonch y Belchi (1987-1990). La evaluación del Potencial de Aprendizaje es la determinación de la capacidad de un sujeto para aprovecharse de un entrenamiento dirigido a mejorar la inteligencia o, lo que es lo mismo, a enseñar a resolver problemas complejos en los que, en principio, parecía fallar (Calero, 1995b; Fernández-Ballesteros, Campllonch, Calero y Belchi, 2000). Con estas premisas, en la actualidad se están desarrollando numerosos programas de entrenamiento cognitivo (ver Calero, 1995c; Calero y García, 1995), así como de entrenamiento en habilidades interpersonales (ver Pelechano, 1995).

En parte, el enfoque del procesamiento de la información ha promovido el desarrollo de una línea alternativa de evaluación del potencial de aprendizaje o evaluación dinámica de la inteligencia versus la evaluación estática de los tests de inteligencia tradicionales. La evaluación dinámica en opinión de Bolig y Day (1993) permite identificar habilidades de aprendizaje de los niños, determinar cómo y qué enseñar, controlar las diferencias individuales... Además, la evaluación del Potencial de Aprendizaje está cerca de la llamada evaluación formativa, que compromete tanto al evaluado como al evaluador en el avance cognitivo y en la adquisición de la madurez personal (Forns y Amador, 1995c).

3.6.2. LA EVALUACIÓN NEUROPSICOLÓGICA

Orígenes del modelo

Como ya había señalado previamente al analizar las fuentes de la evaluación psicológica, los antecedentes más remotos de la evaluación neuropsicológica son los trabajos de Gall (1758-1828). Este médico planteó la existencia de relaciones entre la conducta de sujeto, un área del cerebro responsable de la misma, y su manifestación externa en la forma craneal. Empleó una metodología denominada craneoscopia consistente en realizar análisis de las protuberancias craneales que dieron pie a los primeros mapas del cráneo.

Esta perspectiva localizacionista ha sido replicada posteriormente debido a los resultados de trabajos relacionados con el estudio y tratamiento de sujetos con lesiones cerebrales, ya que

estos estudios constatan la capacidad del cerebro para suplir las funciones asumidas por un área cerebral cuando se produce un daño cerebral en la misma (equipotencialidad).

Esta disciplina toma cuerpo en la década de los años 40, desarrollándose progresivamente hasta la actualidad. El creciente desarrollo de la misma en los últimos años se debe al desarrollo simultáneo de tres líneas de investigación: 1) desarrollo y perfeccionamiento de técnicas neuropsicológicas cada vez más precisas y potentes; 2) grandes avances en la Neurociencia; y 3) las aportaciones de las ciencias de la conducta especialmente de la Psicología Cognitiva.

Presupuestos teórico-conceptuales básicos

El debate sobre las relaciones mente-cerebro es complejo y no puede ser despachado en sólo unas líneas o sólo unas páginas, ya que tanto las posturas dualistas como monistas o emergentistas requieren y necesitan explicarse con sólidos argumentos. En cualquier caso nadie es capaz de negar en la actualidad que un daño, una lesión en el cerebro, puede provocar alteraciones y trastornos psicológicos importantes que pueden afectar tanto a las funciones psicológicas básicas (atención, percepción), a las superiores (razonamiento, memoria, resolución de problemas, lenguaje, imágenes mentales) o a aspectos emocionales, de personalidad... (Fernández-Ballesteros y León Carrión, 1992). Forns et al. (1990) siguiendo a Luria, exponen los presupuestos conceptuales más relevantes sobre los que se apoya la evaluación psiconeurológica:

- 1) Se presupone una interrelación recíproca entre organismo y conducta.
- 2) El concepto de organismo se refiere básicamente al SNC y en especial al cerebro considerando los aspectos anatómico-fisiológicos.
- 3) El cerebro se estructura en tres unidades funcionales básicas (Luria, 1966): (a) la unidad de regulación del tono cortical (ciclo vigilia/sueño, estados atencionales) ubicada en el tallo cerebral, (b) la unidad referida a la adquisición, elaboración, almacenamiento y evocación informativa ubicada en la zona posterior del cortex cerebral; y (c) la unidad que programa, regula y verifica la actividad, ubicada en las zonas anteriores del cortex.
- 4) Cada una de las unidades funcionales se rige por una serie de leyes específicas. La primera unidad regula la actividad vigil del cerebro mediante tres leyes: la ley de la fuerza, la de la concentración y la de la movilidad. La segunda y la tercera unidades funcionales desarrollan su actividad por zonas. En la segunda unidad funcional, las zonas primarias tienen por misión el análisis perceptivo del estímulo, las zonas secundarias son de síntesis perceptual y las zonas terciarias posibilitan una síntesis multimodal de las características del estímulo. En la tercera unidad funcional, las zonas primarias controlan la ejecución motriz de la actividad comportamental, las zonas secundarias programan las secuencias motoras y las terciarias contribuyen a la creación de planes e intenciones y verificación de los resultados.
- 5) Estas unidades funcionales no actúan de forma aislada, por el contrario, existen una serie de principios generales de funcionamiento del cerebro humano (Luria, 1966): (a) el principio de estructuración jerárquica de los tres bloques cerebrales de tal modo que una lesión en el tercer bloque no perturba necesariamente el funcionamiento del primero y el segundo, aunque, si el primer bloque se encuentra afectado, también afectará a los otros dos; (b) el principio de especificidad decreciente de las distintas zonas, mediante el cual las células ubicadas en las zonas primarias, secundarias y terciarias se caracterizan por presentar

distintos niveles de activación según las características modales del estímulo; y c) el principio de lateralización progresiva de las funciones, mediante el cual determinadas funciones tienden a tener una mayor proyección en el hemisferio izquierdo que en el derecho.

Estos preceptos expuestos hacen referencia al modelo de Luria sobre el funcionamiento cerebral, y si bien son tendencialmente aceptados por los neuropsicólogos actuales, algunos de estos puntos han sido contestados. Uno de los temas más controvertidos se refiere a la localización cerebral de los procesos psíquicos superiores. En esta segunda mitad del siglo XX se produjo un resurgimiento de las posturas localizacionistas, si bien desde una perspectiva renovada y crítica. No se aceptó una visión esquemática y reductista de las relaciones entre funcionalismo cerebral y conducta. Se trata de analizar sistemas complejos, áreas cerebrales con funciones específicas y no específicas que establecen entre sí múltiples interacciones y en los que la información recogida del medio, se transforma y se integra progresivamente posibilitando su interpretación cortical lo que da lugar a una conducta adaptada.

Objetivos de evaluación y unidades de análisis

La evaluación neuropsicológica estudia las relaciones funcionales entre conducta y cerebro sustentándose en lo neurofisiológico y desde un punto de vista psicométrico y objetivo. La primaria finalidad de la evaluación neuropsicológica es plantear inferencias acerca de las características estructurales y funcionales del cerebro de la persona, evaluando la conducta de los individuos en definidas situaciones de estímulo-respuesta (Benton, 1994). Para Jiménez y Rodríguez (1992) el objetivo de esta evaluación es diagnosticar la presencia de un daño o disfunción cerebral y, si fuera posible, localizarlo... para facilitar la atención al paciente y su rehabilitación.

Benedet (1986) y Snow y Hooper (1994) incluyen como objetivo evaluador la determinación de funciones preservadas en vista a la planificación de programas de intervención. La evaluación neuropsicológica no se limita al análisis de funciones patológicas y a su tratamiento, sino que intenta dar una explicación del comportamiento basado en el conocimiento de las características funcionales del Sistema Nervioso. Su interés cada vez más destacado por el funcionamiento normal (procesos cognitivos) y por incluir aspectos evolutivos e involutivos ha ampliado su inicial campo de abordaje en relación a las lesiones cerebrales y su tratamiento. Un aspecto esencial de esta disciplina es la de no restringir su ámbito de análisis a los aspectos patológicos del comportamiento, vía de análisis inicial, y su interés cada vez más destacado por el funcionamiento normal (procesos cognitivos) y por incluir aspectos evolutivos dentro de su campo de investigación.

La Neuropsicología intenta describir, identificar y establecer las relaciones existentes entre la organización cerebral y las actividades cognitivas; es decir, el conjunto de procesos que nos permiten el acceso al conocimiento, el aprendizaje y la adaptación al mundo exterior: la percepción, la acción, el lenguaje, la memoria y las actividades de tipo operatorio (Hecaen, 1982). Desde este enfoque de la evaluación se pretende explicar la conducta en sus relaciones con el cerebro y predecirla. Las investigaciones experimentales y clínicas sobre la recuperación funcional responden a este objetivo predictivo. Además de comprender y predecir, la conducta la evaluación neuropsicológica, en la actualidad, tiene por objetivo tratar de establecer las

relaciones entre el funcionamiento cerebral y la naturaleza de los mecanismos cognitivos de procesamiento de la información. El objetivo actual de la Neuropsicología Cognitiva es la comprensión de los diferentes estilos de procesamiento de información que realizan los sujetos y la relación entre tales estilos y las estructuras y funciones cerebrales (Forns et al. 1990).

Aspectos metodológicos y técnicas empleadas

Esta disciplina se caracteriza por su riqueza metodológica y tecnológica. Aborda muy diversos problemas específicos, y por lo tanto, utiliza diversos procedimientos en función del problema objeto de estudio. Metodológicamente emplean tanto procedimientos clínicos como diseños de investigación cercanos a la metodología experimental.

En cuanto a las técnicas son numerosas y se circunscriben al método que se pretende utilizar y a la propia finalidad que persigue la evaluación neuropsicológica. Se han realizado diversas agrupaciones de las mismas (Benton, 1994; Fernández-Ballesteros y León Carrión, 1992; Godoy, 1992; Jiménez y Rodríguez, 1992; Vila, 2004). Integrando las distintas perspectivas de estas referencias podemos clasificarlas en 4 grupos:

1. Técnicas neurológicas importadas de la medicina:

- Tomografía axial computerizada (TAC).
- Resonancia magnética nuclear (RMN).
- Tomografía por emisión de positrones (TEP).
- Determinación del flujo sanguíneo cerebral.
- El electroencefalograma (EEG).
- Potenciales evocados (PE).

2. Clásicos tests o baterías neuropsicológicas propiamente dichas:

- Batería Neuropsicológica Halstead-Reitan (Halstead, 1947; Reitan, 1975).
- El examen Neuropsicológico de Luria-Christensen (Christensen, 1978).
- Test de Copia de una Figura Compleja (Rey, 1959/1987).
- Programa integrado de Exploración Neuropsicológica. Test Barcelona (PIEN) (Peña Casanova, 1987, 1991).

3. Técnicas neuropsicológicas vinculadas al estudio de los procesos perceptivos de la información:

- Técnicas de escucha dicótica (audición).
- Técnicas de visión taquistoscópica.
- Percepción diáptica.

4. Técnicas psicológicas centradas en el análisis de procesos cognitivos, memorísticos, atencionales, perceptivos:

- La escala de Memoria de Wechsler (1945) y WAIS-R NI (Kaplan et al., 1996).
- Test de Retención Visual de Benton (1965, 1974).
- Test Guestáltico Visomotor de Bender (1938/1975)

Benton (1994) clasifica los procedimientos de evaluación en los siguientes apartados: (1) Baterías de tests estándar de evaluación neuropsicológica; (2) Tests de evaluación de las capacidades verbales y la afasia; (3) Tests de evaluación de las capacidades visoperceptuales; (4) Procedimientos para la evaluación de la audición; (5) Instrumentos para la evaluación de sensaciones asociadas al contacto corporal como percepción táctil, kinestésica... (somesthesia); (6) procedimientos de evaluación de las habilidades motoras y práxicas; (7) Tests de evaluación de aprendizaje y memoria; así como (8) Procedimientos de evaluación de las funciones ejecutivas (juicios, flexibilidad mental, toma de decisiones, autoeficacia, planificación...) y el razonamiento abstracto. Por otro lado, Benton (1994) incluye como técnicas a las escalas de puntuación basadas en observaciones de la conducta del paciente en el curso del examen, en el hospital o en la vida diaria, realizadas por el examinador, el personal de hospital o los familiares respectivamente.

Respecto a los requisitos necesarios para realizar una adecuada evaluación neuropsicológica, especialmente con niños y adolescentes, León-Carrión (1995) subraya que esta tarea, es una tarea compleja que requiere un buen adiestramiento tanto en clínica como en técnicas de evaluación psicológica en general y neuropsicológica en particular..., así mismo, requiere conocimiento de las alteraciones neurocognitivas posibles... y conocimientos necesarios para realizar el diseño de programas de rehabilitación neuropsicológica personalizada.

Ámbitos de aplicación

En los primeros momentos la evaluación neuropsicológica surgió en el ámbito médico-patológico hospitalario desde donde se abordaban el estudio de los fenómenos de la alteración y recuperación funcional en el caso de lesiones en el Sistema Nervioso debidos a traumatismos, tumores, o enfermedades cerebrales... En la actualidad, los usos clínicos más frecuentes de la Evaluación Neuropsicológica en la Psicología Clínica y de la Salud son los siguientes: (1) Investigación de la función cerebral; (2) Diagnóstico neuropsicológico para identificación de la lesión o disfunción cerebral; (3) Consecuencias comportamentales de la lesión; y (4) Diseño y evaluación de programas de tratamiento y rehabilitación de alteraciones neuropsicológicas.

Posteriormente, se ha abierto una línea de investigación evolutiva que se lleva a cabo en el ámbito educativo y se centra en el análisis de las relaciones existentes entre los procesos neuropsicológicos y los aprendizajes escolares. Así mismo, y en el ámbito de la investigación básica, la neuropsicología cognitiva está estudiando las relaciones de los procesos cognitivos superiores con su sustrato neurológico.

3.7. RESUMEN: CARACTERÍSTICAS DIFERENCIALES DE LOS MODELOS EN LA EVALUACIÓN PSICOLÓGICA

En este apartado y, a modo de síntesis, se presenta el Cuadro 8 en el que se exponen algunos referentes significativos de cada modelo en relación a sus orígenes, presupuestos teóricos básicos, variables objeto de estudio y objetivos de la evaluación, así como los métodos-técnicas empleadas y sus ámbitos de aplicación.

Cuadro 8. Características diferenciales de los modelos en la Evaluación Psicológica

MODELO PSICOMÉTRICO CORRELACIONAL O DEL ATRIBUTO	
Orígenes	Galton, Mckeen Catell, Binet. Otros representantes: Eysenck, Guilford.
Presupuestos Teóricos	- La conducta se explica en función de variables internas o intrapsíquicas (rasgos). - Estas variables dan consistencia y estabilidad al comportamiento a lo largo del tiempo y de las situaciones. - Rasgo: innato. - Reformulaciones posteriores: persona x situación (interaccionismo)
Variables	Rasgos de la personalidad, factores, dimensiones, características
Objetivos	- Describir, Clasificar-Comparar, y Predecir. - Estudiar las diferencias individuales y ubicar al sujeto en relación al grupo de referencia normativo en un rasgo o dimensión concreta que puede cuantificarse. - Enfoque nomotético.
Métodos	Método Correlacional.
Técnicas	Técnicas psicométricas, elaboradas mediante procedimientos estadísticos y factoriales: Cuestionarios de Personalidad, Tests de inteligencia, Tests de intereses...
Ámbitos de aplicación	Escolar. Organizacional. Clínico. Investigación

MODELO CLÍNICO-DINÁMICO: PERSPECTIVA MÉDICO-PSIQUIÁTRICA	
Orígenes	Incremento de las necesidades de evaluación clínica por efecto de las dos guerras mundiales.
Presupuestos Teóricos	- La conducta esta determinada por causas genético-biológicas e intrapsíquicas. - Modelo médico de enfermedad.
Variables	- Entidades nosológicas. - Trastornos neurobiológicos.
Objetivos	- Describir, Clasificar, Predecir y Pronosticar. Explicar y Rehabilitar. - Estudian el grado de afectación del trastorno, los factores histórico-biográficos etiológicos, el curso del proceso psicopatológico y el tratamiento farmacológico adecuado.
Métodos	Método Observacional, Correlacional y Experimental
Técnicas	Tests psicométricos para constatar rasgos, dimensiones o entidades nosológicas; Técnicas observacionales; Técnicas proyectivas. Escalas de observación.
Ámbitos de aplicación	Clínico.

MODELO CLÍNICO-DINÁMICO: PERSPECTIVA PSICOANALÍTICA	
Orígenes	Modelo Freudiano de la Personalidad.
Presupuestos Teóricos	- La conducta se explica en base a una serie de construcciones teóricas internas que conforman la estructura de la personalidad en conexión con una serie de dinamismos internos inconscientes. - Énfasis en la historia pasada del sujeto. - Estudian las elaboraciones mentales subjetivas del sujeto, más que su conducta manifiesta.
Variables	- Estructura intrapsíquica de la personalidad del sujeto (Ello/Yo/Super-Yo). - Peso de los elementos topológicos (Consciente, Preconsciente, Inconsciente). - Mecanismos de defensa.
Objetivos	Describir, Clasificar, Explicar, Comprender y Predecir.
Métodos	Observación Clínica y Estrategias correlacionales.
Técnicas	Técnicas proyectivas (preferentemente); Análisis de los Sueños; Asociación Libre; Entrevista libre o semiestructurada; Autobiografías; Análisis de la Transferencia. Técnicas de enfoque idiográfico, aunque pueden usar técnicas que posean una vertiente nomotética.
Ámbitos de aplicación	Clínico (preferentemente) y Escolar.

MODELO FENOMENOLÓGICO, HUMANISTA, EXISTENCIAL - CONSTRUCTIVISTA	
Orígenes	Husserl, Rogers, Maslow, Perls...
Presupuestos Teóricos	La conducta está determinada por la percepción subjetiva que el sujeto tiene del mundo y de sí mismo, por los estilos comunicacionales, y por el potencial humano (tendencia a la autoactualización).
Variables	Autopercepción, Percepción de los otros y del ambiente, Estrategias personales de resolución de problemas, Tendencia a la autoactualización, Motivación de crecimiento... Conocimiento construido. Constructos personales. Planes y Significados.
Objetivos	Comprender e Intervenir para estimular el crecimiento personal, logrando un mejor ajuste personal y social del sujeto.
Métodos	Observación Clínica y Estrategias correlacionales.
Técnicas	Técnicas subjetivas (preferentemente); Observación y Autoobservación; Entrevista libre; Técnicas proyectivas; Técnicas psicométricas; Autobiografía; Documentos personales.
Ámbitos de aplicación	Clínico (preferentemente) y Escolar.

MODELO CONDUCTUAL RADICAL (PRIMERA GENERACIÓN)	
Orígenes	Watson, Skinner...
Presupuestos teóricos	• Condicionamiento clásico y operante. • El ambiente como determinante directo y único de la conducta del sujeto, por lo que el tratamiento se dirige a la modificación de variables ambientales. • Rechaza la introspección, proponiendo una explicación de la conducta en términos estímulo-respuesta, respuesta-reforzamiento.
Variables	• Análisis de las conductas directamente observables (Análisis topográfico de la conducta). • Variables ambientales y motoras.
Objetivos	• Controlar y Predecir la conducta.
Métodos	Metodología Observacional y Experimental.
Técnicas	Técnicas observacionales: técnicas con garantías psicométricas de validez y fiabilidad, así como con precisión y sensibilidad para detectar: la ocurrencia de la conducta, la repetibilidad, la ocurrencia en distintas condiciones estímulares, con parámetros que mejor definan la conducta y la relación entre la conducta y otras conductas del sujeto.
Ámbitos de aplicación	Laboratorio (inicial y preferentemente), Clínica, Escolar....

MODELO CONDUCTUAL MEDIACIONAL O NEOCONDUCTISMO (SEGUNDA GENERACIÓN)	
Orígenes	Tolman, Guthrie, Hull,..... Wolpe, Eysenck, Bandura, Rotter...
Presupuestos teóricos	• El ambiente tiene gran importancia en la determinación de la conducta, pero el impacto de la situación sobre la conducta está mediatizado por una serie de variables del organismo o variables intervinientes. • Interés por los procesos mediacionales (perceptos, imágenes, ideas...), es decir, por los procesos mentales que pueden ser estímulos encubiertos.
Variables	• Aceptan la medida de variables no directamente observables. • Variables mediacionales diversas de carácter cognitivo-emocional: Intención, Ansiedad, Motivación, Procesos cognitivos intermediarios (valor del refuerzo para el sujeto, locus de control...).
Objetivos	Predecir y Controlar la conducta.
Métodos	Método experimental.
Técnicas	Autoinformes Verbales y Cuestionarios sobre conducta cognitiva; Técnicas de observación; Técnicas de intervención en la ansiedad (modelado, desensibilización sistemática, relajación...).
Ámbitos de aplicación	Clínico (preferentemente), Escolar, Comunitario.

MODELO CONDUCTUAL INTEGRADOR-INTERACTIVO (TERCERA GENERACIÓN): LA EVALUACIÓN CONDUCTUAL-COGNITIVA	
Orígenes	Ellis, Cautela, Mahoney, Beck, Meichenbaum
Presupuestos teóricos	Incluyendo al ambiente como variable elicitor de conductas, estiman que su impacto sobre el organismo no es directo ni siquiera mediatizado, sino interactuante con las propias variables de cada organismo.
Variables	- Analizan variables ambientales, variables del organismo y variables de respuesta. - Tres niveles en la evaluación del organismo: <i>Conductas cognitivas</i> (los pensamientos, la interpretación del medio que efectúa el sujeto, las expectativas y los automensajes, habilidades o estrategias, motivaciones...), <i>Conductas motoras</i> y <i>Conductas fisiológicas</i>.
Objetivos	Análisis topográfico, Explicación Funcional y Modificación de la conducta. Explicar, Controlar y Predecir la Conducta.
Métodos	Metodología Observacional, Experimental y Correlacional.
Técnicas	Observación-Autoobservación; Autoinformes; Registros psicofisiológicos.
Ámbitos de aplicación	Clínico (preferentemente), Escolar, Organizacional.

MODELO CONDUCTUAL INTEGRADOR-INTERACTIVO (TERCERA GENERACIÓN): EL CONDUCTISMO PARADIGMÁTICO	
Orígenes	Staats.
Presupuestos teóricos	- No acepta la linealidad en los determinantes de la conducta e incluyen en su formulación variables situacionales, variables de persona, de respuesta y variables biológicas. - Tiende un puente entre el conocimiento conductista y el tradicional.
Variables	- Condiciones ambientales históricas de aprendizaje del sujeto y Condiciones ambientales actuales. - Repertorios Básicos de Conducta (elementos que configuran la personalidad): Repertorio Lingüístico-Cognitivo, Sensorio-Motor; y Emocional-Motivacional. - Condiciones biológicas pasadas y actuales. - Topografía de la conducta problema.
Objetivos	Explicación y Modificación de la Conducta.
Métodos	Método Naturalista-Experimental (estudia la conducta en ambiente natural, introduce variaciones y controles para conocerla con precisión).
Técnicas	Tests psicométricos de inteligencia y personalidad tradicionales (interpretados como muestra de lo que el sujeto es capaz de hacer a determinada edad); Autoinformes escritos y verbales (sobre emociones y pensamientos); Registros psicofisiológicos; Técnicas observacionales.
Ámbitos de aplicación	Clínico (preferentemente).

MODELO COGNITIVO	
Orígenes	Influencias: Avances de la electrónica y con ella el impacto del ordenador, Nuevas Formulaciones psicolingüísticas (Chomsky), La obra de Piaget, El enfoque de la interacción biológico-social soviético....
Presupuestos Teóricos	• El sujeto es un agente activo, procesador de información del medio externo e interno. • La conducta es explicada en base a una serie de procesos y estructuras mentales internas, es una función del mundo cognitivo de la persona. • No se niega la importancia del aprendizaje, pero se admite cierta preprogramación y potencial biológico. • Destacan cierta independencia del funcionamiento cognitivo con respecto a las variables del ambiente, considerando que el sujeto cuenta con planes, propósitos, programas de acción que excede a la conducta elicitada.
Variables	Estudian variables cognitivas: Estructuras, Procesos, Estrategias cognitivas... Cogniciones, atribuciones, funciones mentales superiores como la memoria o el lenguaje, representaciones, fases en el procesamiento de la información, procesos y estrategias cognitivas empleadas para resolver tareas cognitivas, relación entre operaciones cognitivas y variables fisiológicas, relación entre diferentes estilos cognitivos y diferencias individuales....
Objetivos	Describir, Explicar y Predecir la conducta.
Métodos	Metodología Experimental, Estrategias Correlacionales y Metodología Observacional.
Técnicas	Autoinformes; Tests situacionales (simulación); Instrumentos de evaluación fisiológica (procedimiento cronométricos...); Baterías Neuropsicológicas; Pensamientos en voz alta; Ejecuciones ante tareas cognitivas.
Ámbitos de aplicación	Laboratorio, Escolar y Clínico (neurociencias....).

EVALUACIÓN DESDE EL PROCESAMIENTO DE LA INFORMACIÓN	
Orígenes	Fusión de la Psicología experimental con la Teoría y tecnología computacional (inteligencia artificial).
Presupuestos Teóricos	• Analogía entre procesos computacionales y conceptos mentales, pero la igualdad no es virtual, ya que en el ser humano se da procesamiento lineal como en las máquinas, y otro tipo de procesamiento simultáneo o paralelo basado en el supuesto de que el sistema cognitivo humano posee una serie de estadios secuenciales para procesar la información. • Sujeto: Procesador activo de información procedente tanto del medio externo como del interno. • Las capacidades innatas y los procesos internos como planes, estrategias, decisiones... toman relieve. • La organización de la cognición descansa en la articulación dinámica de estructuras cognitivas (redes, esquemas... con sustrato neurológico y cierta estabilidad) y procesos cognitivos.
Variables	Estructuras cognitivas y procesos cognitivos
Objetivos	• Identificación de las estructuras y procesos cognitivos. • Determinación de las variables que influyen en unas y otros y sus complejas interrelaciones.
Métodos	Metodología experimental. Introspección para acceder a lo cognitivo pero objetivada, medible y controlable. Desarrollo de otros métodos que hacen viable el análisis de datos cualitativos.
Técnicas	Análisis cronométricos, programas de ordenador, diagramas de flujo, técnicas de medida de tiempos de reacción, análisis de los movimientos oculares, mapeos cognitivos, procedimientos de medida de la memoria, técnicas taquiscópicas...
Ámbitos de aplicación	Clínico. Escolar.

EVALUACIÓN NEUROPSICOLÓGICA	
Orígenes	Gall. Representantes: Benedet, Benton...
Presupuestos Teóricos	Existencia de relaciones entre el cerebro (SNC) y las funciones psicológicas básicas (atención, percepción,...) las superiores (razonamiento, memoria, resolución de problemas, lenguaje, imágenes mentales..) y aspectos emocionales, de personalidad...; Relaciones entre funcionamiento cerebral y conducta.
Variables	Funciones psicológicas básicas y superiores. Estructuras y procesos cognitivos. Funcionamiento cerebral.
Objetivos	• Estudiar las relaciones funcionales entre conducta y cerebro sustentándose en lo neurofisiológico y desde un punto de vista psicométrico y objetivo. • Describir, identificar y establecer relaciones entre la organización cerebral y las actividades cognitivas, afin de explicar la conducta y predecirla.
Métodos	Metodología experimental. Procedimientos Clínicos.
Técnicas	• Técnicas neurológicas importadas de la medicina: TAC, RMN, EEG, PE... • Clásicos test o baterías neuropsicológicas: Halstead-Reitan, Luria-Christensen, Test de copia de una figura compleja de Rey, Test de Barcelona... • Técnicas neuropsicológicas vinculadas al estudio de los procesos perceptivos de la información: Escucha dicótica, visión taquistoscópica,... • Técnicas psicológicas centradas en el análisis de los procesos cognitivos, memorísticos, atencionales, perceptivos...
Ámbitos de aplicación	Clínico. Educativo. Investigación básica.

CAPÍTULO 4.
MÉTODOS Y CONTROVERSIAS EN EVALUACIÓN PSICOLÓGICA

- 4.1. Introducción
- 4.2. Principales métodos en la evaluación psicológica
 - 4.2.1. Método Correlacional
 - 4.2.2. Método Experimental
 - 4.2.3. Método Observacional
- 4.3. Alternativas metodológicas polémicas
 - 4.3.1. Generalidad - Especificidad de la conducta
 - 4.3.2. Idiográfico - Nomotético
 - 4.3.3. Cuantitativo - Cualitativo
 - 4.3.4. Explicación – Comprensión
- 4.4. Resumen: Principales métodos y alternativas polémicas

4.1. INTRODUCCIÓN

Como ya se ha señalado previamente, la evaluación psicológica tiene entre sus tareas determinar qué métodos son más idóneos para optimizar la viabilidad de sus objetivos con las máximas garantías científicas. Por ello, y con la finalidad de clarificar diversas cuestiones metodológicas de esta disciplina, en este capítulo se abordarán algunos aspectos en relación a los principales métodos utilizados en Evaluación Psicológica, así como diversas alternativas metodológicas polémicas. En primer lugar, se describen los tres métodos básicos, el correlacional, el experimental, y el observacional, en sus características y aplicaciones más específicas. Posteriormente se plantean diversas polémicas derivadas de los distintos modelos y métodos empleados en evaluación psicológica, polémicas que se articulan en torno a 4 ejes: cuantitativo-cualitativo, idiográfico-nomotético, generalidad-especificidad de la conducta, y explicación-comprensión de la misma.

4.2. PRINCIPALES MÉTODOS EN LA EVALUACIÓN PSICOLÓGICA

Suele definirse el método como cualquier forma de proceder para llegar a un fin, siempre y cuando tanto el fin como la forma estén establecidos de antemano y puedan ser utilizados por cualquiera (Tous, 1989a), o la forma de llevar a cabo una acción de manera estructurada. Es decir, el método es el procedimiento sistemático que hace viable un objetivo. Es precisamente este sistematismo, esta estructuración procesal lo que diferencia el método científico del que no lo es, y lo que permite, entre otras, la replicabilidad del conocimiento, indispensable para que la ciencia avance.

Para Yela (1994) el método científico es un camino de acceso a la intelección de la realidad, y consiste en dar razón sistemática, empírica y en lo posible experimental de los fenómenos. El principio básico del método científico natural es posible generalizarlo del siguiente modo: El objeto de estudio, sea el que sea, sólo es posible llegar a conocerlo por el mismo, mediante la observación, y la manipulación y control si es necesario, de los agentes distintos al mismo que nos permiten explicar su propia naturaleza y funcionamiento, a través de los fenómenos que aquellos agentes producen en el mismo de forma observable (Tous, 1988b).

El método consiste en un conjunto de procedimientos sistemáticos y organizados que llevan a dar respuestas a los problemas planteados (Arnau, 1996a). En otras palabras, los hechos y los fenómenos deben ser relacionados entre sí, de una forma lógica y sistemática, para construir un conocimiento organizado, pudiéndose realizar esta relación a tres niveles: descriptivo, relacional o explicativo, por lo tanto, describir, explicar, predecir y controlar se van a constituir en fines de la ciencia (Clark-Carter, 2002; Navas, 2001).

El avance del conocimiento y la comprensión de los fenómenos se alcanza a través de la investigación. La investigación científica trata de simplificar la complejidad de la naturaleza aislando un determinado fenómeno para que sea sometido a un estudio cuidadoso (Kazdin, 2001). La metodología hace referencia a los diferentes principios, procedimientos y prácticas que dirigen la investigación. Por ello, la metodología enseña formas de pensar sobre las relaciones entre variables, de las causas y efectos, y de las conclusiones obtenidas de la teoría, la investigación y la experiencia. Se considera que es una forma de pensar porque orienta las preguntas que se deben hacer y las prácticas que se diseñan para contestar a esas preguntas.

Las distintas metodologías constituyen el puente que une lo conceptual con lo empírico y aportan el fundamento objetivo a los modelos teóricos. Todo investigador para obtener resultados objetivos y concluyentes utiliza métodos estadísticos, plantea sus problemas mediante el método científico (Arnau, 1996a). Dicho método proporciona una forma de actuación a través de operaciones y fases ordenadas que van a ir orientando la investigación a fin de cumplir los fines de la ciencia. Aunque el método científico es un proceso general y unitario, se puede dividir en diferentes fases (Arnau, 1996a; Balluerka y Vergara, 2002):

- Fase teórico-conceptual: Identificación del problema y formulación de hipótesis que se puedan contrastar empíricamente.
- Aplicación técnico-metodológica: Operativización de las hipótesis, elaboración del plan de investigación y selección de una estrategia de obtención de datos.
- Fase estadístico-analítica: Contrastación de las hipótesis propuestas, interpretación de datos, conclusiones, retos futuros y elaboración del informe.

La evaluación psicológica en tanto que disciplina científica, se adhiere al método general de la ciencia y, más concretamente, al método hipotético-deductivo, considerado como el método científico positivo por excelencia. Pese a este consenso respecto al método científico general, el problema surge al elegir las estrategias concretas mediante las cuales hacer viable este método general. Así, se han perfilado tres metodologías: la correlacional, la experimental y la observacional.

La polémica experimental-correlacional tiene un punto álgido en 1957, a raíz del artículo de Cronbach en el que denuncia las dos corrientes en que se había polarizado la psicología científica. La corriente experimental tiene por objetivo estudiar los fenómenos bajo condiciones de estricto control, a fin de alcanzar los principios explicativos generales que los rigen. Busca establecer relaciones de causalidad que le permita imputar a la manipulación de las variables independientes los cambios observados en las variables dependientes. Las diferencias individuales no sólo no han sido un punto de interés, sino un claro estorbo, “siendo arrojadas a estas tinieblas exteriores conocidas con el nombre de varianza de error” (Cronbach, 1957). En contraposición, la corriente correlacional fija su punto de interés justamente en esta supuesta varianza de error, en las diferencias que se observan entre los individuos. El control de las variables se realiza mediante procedimientos estadísticos y no físicos, no existe restricción alguna al tipo de estudio y variables a controlar... y se aplica una lógica de atribución multicausal a los fenómenos. Su método básico de análisis es el correlacional, que busca establecer relaciones de concomitancia entre los hechos tal como se producen naturalmente, sin la manipulación del experimentador. Pese a esta dicotomía metódica es evidente que la Psicología se nutre tanto de los estudios provenientes de la psicología experimental como de la correlacional, y así mismo de la metodología observacional. Además, la posibilidad de combinar enfoques sin perjuicio del status científico de la psicología, está claramente explicitada en la obra de muchos investigadores.

4.2.1. EL MÉTODO CORRELACIONAL

La psicología correlacional se puede llegar a definir como la rama de la metodología psicológica que intenta encontrar relaciones entre variables, que ya se encuentran dadas en la naturaleza. Su finalidad es estimar el grado o fuerza de covariación entre las mismas.

El término correlación hace referencia a la medición estadística de la relación recíproca entre variables, indicando la dirección y magnitud de dicha relación. Los métodos correlacionales, mediante el establecimiento del grado de correlación entre variables, posibilitan que se realice una predicción de los resultados que se pueden encontrar en una variable (criterio) a partir de los resultados obtenidos en otra (predictor). Ahora bien, es importante señalar que esto no significa que el predictor se pueda tomar como variable causal, ya que pueden existir otras variables responsables de la correlación.

El objetivo de la psicología correlacional ha sido la evaluación de atributos, de cara a encontrar las diferencias individuales, y clasificar a los sujetos, tanto en lo referente a funciones cognitivas como de personalidad. En este caso, se trabaja con el supuesto de la estabilidad conductual, se trata del diagnóstico de estado, orientado a normas, que utiliza tests o instrumentos que ofrecen una determinada puntuación, acerca de determinado rasgo, con el fin de determinar su presencia relativa en un sujeto.

El método correlacional es aplicable a aquellas situaciones en que se busca establecer relaciones de concomitancia entre los hechos tal como se producen naturalmente, sin la manipulación del experimentador y se pretende un objetivo de generalización y predicción. Se trata de una metodología centrada en hallar las diferencias individuales en aspectos consistentes del comportamiento humano, y desde este método se han intentado, no sólo determinar la fiabilidad y validez de las técnicas, sino estudiar también la estructura de la inteligencia y de la personalidad.

El Psicodiagnóstico ha estado muy vinculado desde sus inicios a este tipo de método, siendo un claro exponente de ello el modelo de rasgos y factores. El psicodiagnóstico nace y se desarrolla unido al modelo correlacional, o más bien dicho, constituye su cara aplicada, de la que la cara más teórica está relacionada con los enfoques o disciplinas psicológico-diferenciales.

Las críticas que han recibido estos métodos se referirán, principalmente, a las técnicas que se han utilizado y a las conclusiones, debido a que en ocasiones se ha llegado al establecimiento de inferencias causales a partir de la obtención, únicamente, de correlación entre las variables objeto de estudio. A pesar de ello, estos estudios son útiles en las primeras etapas de la investigación diagnóstica porque facilitan la descripción de relaciones.

Este método dirigido primordialmente al estudio de la estructura de la inteligencia y de la personalidad ha obligado a establecer conceptos clave en la investigación estadística, tales como fiabilidad, validez o varianza de error. Además, en los últimos años se emplean técnicas complejas (regresión múltiple, análisis discriminante...) en un intento de complementar estudios correlacionales con predictivos. Aunque actualmente los procedimientos multivariados (aplicación de un análisis factorial a una matriz de datos) tienen gran vigor en el estudio de las dependencias e interdependencias entre variables, se sigue sin poder establecer una relación de causalidad entre las mismas. El hecho de que el análisis multivariante sea el más utilizado deriva de la ventaja de permitir estudiar las interacciones existentes entre variables. Así, conviene tener presente que las técnicas asociadas al método correlacional se han visto

enriquecidas por la utilización de complejas técnicas multivariantes y por el uso de los ordenadores.

Los diseños correlacionales según la finalidad de la investigación se han dividido en (Martínez Arias, 1987a):

- *Diseños descriptivos o de covariación.* Con una finalidad esencialmente descriptiva analizan las variaciones que se producen entre diferentes variables a través del análisis de cluster o del análisis factorial.
- *Diseños predictivos.* La predicción consiste en estimar posibles valores de una variable dependiente o criterio a partir de los que presenta otras independientes o predictoras, para ello se parte de las relaciones previamente identificadas entre variables predictoras y criterios. Estos diseños se basan en la ecuación de regresión aplicándose técnicas como la regresión múltiple y el análisis discriminante.
- *Diseños correlacionales causales.* El objetivo que se persigue con los mismos es conocer los factores asociados con ciertos hechos mediante la explicación de las relaciones causales existentes entre diferentes variables. A la matriz de correlaciones se aplican técnicas como el path análisis, que si bien no permite realizar un análisis causal propiamente dicho, permite comprobar si ciertos modelos causales son matemáticamente apropiados.

Las diferencias entre el método correlacional y experimental definieron las dos disciplinas metodológicas de mayor tradición. Sin embargo, progresivamente se ha destacado la necesidad de emplear ambos enfoques para investigar tanto las diferencias individuales como los efectos del tratamiento. De hecho algunos estudios (Kendall y Norton-Ford, 1988) inciden en la importancia de planificar investigaciones que estudien la interacción entre las diferencias individuales y los efectos de los tratamientos.

Silva (1985) dedica en su obra un apartado titulado “el psicodiagnóstico entre la perspectiva correlacional y experimental”, concluyendo que, ya se trate de un diagnóstico de base correlacional o de base experimental, se encuentra con los problemas que producen y llevan a su vez a la consideración de efectos moduladores (correlacional) o de efectos de interacción (experimental). Sea uno u otro caso, se aconseja lo mismo: un tratamiento lo más individualizado posible del problema que tenemos delante, un enfoque máximamente diferenciador (en los límites de la utilidad), un enfoque situativo que tome en cuenta el máximo número de variables tanto del sujeto como del contexto, la interacción entre ambos y el carácter esencialmente activo del sujeto en la determinación del medio. Sólo un enfoque tal asegurará una mejor calidad.

La predominancia de un método y otro dentro de un área de estudio concreta es, por lo tanto, coyuntural y no puede implicar una adscripción definitiva. Es mucho más realista considerar los dos métodos como aproximaciones diferentes, con sus propias ventajas y limitaciones, lo que hace que valgan para abordar algunos tipos de problemas pero no otros. Cuando el tipo de problemas que constituyen el objeto de un área de la psicología cambia, es probable, que uno de los dos métodos ceda la prioridad al otro. Esto no quiere decir que no se deba intentar una aproximación entre los dos métodos como sugiere Cronbach.

4.2.2. EL MÉTODO EXPERIMENTAL

Procede del enfoque galiléico y su metodología de trabajo se caracteriza por ser hipotética-deductiva, al igual que la Física, ciencia que tanta influencia ha tenido en la Psicología. El método experimental ha sido identificado con el método científico por su pretensión de hallar una explicación causal de los fenómenos observados. Comprende una obtención de datos, el análisis correspondiente de los mismos, e implica que la investigación sea observable, controlada, empírica y repetible.

La metodología experimental persigue como objetivo estimar la magnitud del efecto de las variables de tratamiento o variables independientes o causales (Balluerka y Vergara, 2002) y se fundamenta en tres pilares: aleatorización, control e hipótesis de causalidad (Gad, 1999). El control, la manipulación experimental y la aleatorización en la selección de la muestra y asignación de los sujetos a grupos son las características esenciales de este método (Balluerka y Vergara, 2002). El método experimental es el método que mejor cumple las exigencias del método científico ya que trata de llegar a conocer si una determinada variable, sometida a manipulación, es la causa de los cambios que se observan en otra, así que este método se caracteriza por la búsqueda de relaciones causales entre variables. Supone una obtención sistemática de datos, el análisis de los mismos, y que la investigación sea observable, controlada, empírica y repetible. Por lo que respecta a nuestra disciplina hay que señalar que las técnicas de evaluación son utilizadas generalmente como variables dependientes y, a veces, como independientes en los diseños experimentales. En este método el experimentador elabora una hipótesis en la que presupone que una variable (variable independiente) es la causa de otra que depende de la anterior (variable dependiente). A fin de comprobar este supuesto, el investigador efectúa una manipulación sistemática sobre la variable independiente para observar y analizar su efecto en la variable dependiente.

El interés de la Psicología por conocer la relación causal entre determinadas variables ha dado lugar a la utilización creciente de este método. En este enfoque el experimentador elabora una hipótesis en la que presupone que X es la causa de Y. Para su comprobación actúa sistemáticamente sobre la variable X con el propósito de analizar su efecto en la variable Y. La manipulación de la variable que se hipotetiza como causa se denomina variable independiente (VI). La conducta o cambios en la conducta que se observan como efecto de lo anterior, se denomina variable dependiente (VD). Las demás variables deben estar controladas o permanecer constantes para asegurar un buen diseño experimental.

El rigor metodológico exige una situación de control, es decir, un sujeto o grupo de sujetos que no sea sometido a la manipulación (grupo de control) a fin de comparar los cambios con el grupo de sujetos experimentales, y lograr una atribución causal. Así mismo, en un diseño experimental se controlan las variables extrañas que pueden afectar tales como: variables de la situación experimental, las características de la demanda, variables personales y ambientales... Generalmente, este tipo de metodología se emplea en situaciones de laboratorio, aunque también existen diseños de campo o diseños en ambientes ecológicos. Otra variable de control hace referencia a la asignación de los sujetos a investigar para lo que se emplean diversas técnicas tales como asignación aleatoria, apareamiento, la técnica de ciego o doble ciego...

Así, el método experimental permite al investigador concentrarse sobre los posibles efectos en un solo factor o de dos factores (univariados, multivariados), manteniendo constantes los factores que no están sometidos a prueba. Algunos investigadores (León y

Montero, 1996; Weiner, 1976) han enunciado las condiciones que deben darse en un experimento para que pueda realizarse una inferencia causal entre ambas variables:

- Contigüidad temporal relativa, es decir, la variable independiente y la dependiente deben formar secuencia temporal.
- La variable independiente debe preceder a la dependiente, es decir, la causa debe preceder obligatoriamente al efecto.
- El resto de las variables que influyen o pueden influir en la situación experimental deben mantenerse controladas (no expuestas a varianzas).
- En caso de repetirse el experimento deben de producirse los mismos efectos, es decir, los resultados deben ser susceptibles de replicación.

Cuando cumplidas estas condiciones se estima terminado el experimento es posible establecer relación entre los resultados obtenidos y la formulación teórica de la cual se partió. En la medida en que se amplían los experimentos y se ratifican las hipótesis iniciales pueden formularse leyes generales (que quedarán expuestas a ratificación, modificación... en función de los hallazgos que se vayan realizando).

El método experimental ha sido valorado positivamente, pero también ha recibido críticas referidas sobre todo porque se opera en situación de laboratorio y por la dificultad de generalización de sus resultados. Algunos autores plantean que esta metodología trata de comprobar las relaciones de causalidad entre variables, sin embargo, nunca es del todo posible controlar todo menos un aspecto para llegar a la relación causa-efecto. De hecho estos modelos incluyen un término de error.

En la respuesta a estas críticas se argumenta que el experimento debe descomponer “artificialmente” la conducta en unidades parciales como única posibilidad de formular leyes generales. La observación a que es sometida la conducta en una situación experimental no equivale a la que se podría realizar a una persona en la vida diaria; se trata de una observación peculiar, cuya finalidad es evaluar a fin de establecer principios teóricos (Myers, 1988).

El tipo de estrategia metodológica utilizada en un experimento recibe el nombre de diseño experimental. Se han realizado diversas clasificaciones de los diseños experimentales tomando como referencia diferentes criterios.

En la clasificación de Arnau (1996a), se opera una clasificación en función de la cantidad de variables independientes manipuladas, identificando los *diseños de grupo* que se subdividen en dos grandes categorías: *unifactoriales* y *factoriales*.

La forma más simple de diseño de grupo consiste en aplicar tratamientos diferentes a uno o varios grupos de sujetos. A estos diseños en los que se manipula una sola variable para comprobar su efecto se les denomina *diseños unifactoriales de grupo*. Según diversos autores (Gliner y Morgan, 2000; Martínez Arias, 1978b), se pueden distinguir los siguientes diseños unifactoriales: diseño pretest-posttest con grupo control, diseño sólo posttest, diseño Solomon de cuatro grupos, diseño de atención con grupo control y diseño placebo con grupo control.

Cuando se manipulan dos o más variables independientes con al menos dos condiciones experimentales por variable se les denomina *diseños factoriales de grupo*, dichos diseños intentan comparar los efectos simultáneos de dos o más variables independientes. Son diseños que permiten investigar varios factores conjuntamente, así como el efecto de su interacción. Los diseños factoriales son útiles porque permiten averiguar la manera en que las variables

independientes se combinan o actúan recíprocamente para influir sobre la conducta, y porque los efectos de diversas variables pueden examinarse al mismo tiempo en un solo experimento. La estructura del diseño viene determinada por la combinación de la cantidad de variables independientes y la cantidad de valores adoptados por cada variable independiente. En el diseño factorial se puede emplear cualquier número de variables y niveles, pero a mayor cantidad también será necesario mayor número de sujetos. Por lo que los diseños se hacen más complejos a medida que se incrementa el número de variables independientes que se quieran manipular y la diversidad de los efectos conductuales que conlleven (Campbell y Stanley, 1963/1988; Cook y Campbell, 1981).

En función de la estrategia empleada para la comparación entre los tratamientos administrados a los sujetos, se pueden distinguir tres tipos de diseños: los de medidas independientes (intergrupos), los de medidas repetidas (intragrupos) y los de medidas parcialmente repetidas (mixtos). En los *diseños intergrupos* cada tratamiento se administra a un grupo distinto de sujetos, y los efectos de la variable independiente se infieren comparando las medias obtenidas en la/s variable/s dependiente/s en los diferentes grupos (Balluerka y Vergara, 2002). En los *diseños intragrupos* a todos los sujetos se les aplican todas las condiciones experimentales. Los efectos de la variable independiente se estudian al comparar su rendimiento en las diferentes condiciones. El problema que plantean es el denominado *efecto residual, de orden* (Arnau, 1996b) o *de práctica* (León y Montero, 1996), para evitar este problema una de las técnicas a la que se recurre es al *contrabalanceo* que consiste en aleatorizar el orden de aplicación de los tratamientos. En los *diseños mixtos* se combinan las estrategias inter e intragrupos.

Entre las clasificaciones de los diseños experimentales cabe destacar la de Morales Domínguez (1987) que presentan una clasificación tomando como referencia el número de sujetos o de las variables:

Según los Sujetos:

- Un sólo grupo de sujetos: Se efectúan comparaciones entre dos o más tratamientos realizados a un mismo grupo de sujetos aunque desde el punto de vista de la lógica experimental, es como si tuviéramos dos o más grupos distintos. Este tipo de diseño se denomina INTRASUJETOS. Un caso particular de estos, muy discutido desde el punto de vista metodológico, es cuando sólo se utiliza un único sujeto. A este caso se le denomina diseño DE CASO ÚNICO ($N = 1$).
- Más de un grupo de sujetos, denominándose diseños INTERSUJETOS. Estos incluyen: Diseños de dos grupos aleatorios, Diseños de dos grupos equiparados, y Diseños de más de dos grupos.

Según las variables:

- Una sola variable dependiente, es decir, un sólo efecto, que se denomina diseño UNIVARIADO. Cuando en este diseño se estudia una sola causa, es decir, una sola variable independiente se denomina diseño de COMPARACIÓN, mientras que cuando existen más de una variable independiente se denomina diseño FACTORIAL.
- Más de una variable dependiente, es decir, más de un efecto, que se denomina diseño MULTIVARIADO.

Tomando como referencia la mayor parte de las clasificaciones de los diseños experimentales, a continuación clarificaremos las características de los diseños intrasujeto y de los diseños intergrupo. Ambos poseen unos elementos comunes entre sí, esto es: a) se llevan a cabo, en la mayoría de las ocasiones, en laboratorios; b) tienen un carácter hipotético-deductivo por cuanto disponen de una base teórica mediante la cual intentan controlar las variables extrañas; y c) otras características tales como aleatorización, uso de análisis de varianza para el estudio de los resultados...

4.2.2.1. Diseño intrasujeto

El diseño intrasujeto consiste en la aplicación del mismo tratamiento experimental a todos los sujetos de un grupo. En una segunda fase, se comparan entre sí las diferentes puntuaciones obtenidas por cada sujeto (a fin de comprobar la influencia de la variable independiente). Una de las ventajas de este tipo de diseño es la de reducir el error experimental al eliminar la varianza intergrupos. Se trata de un tipo de diseño con el que se economiza tiempo, pues se trabaja con pocos sujetos y no se requiere grupo de control. Por otra parte, dado que permite controlar los efectos que provienen de las diferencias individuales, posibilita valorar los cambios que se producen en cada uno de los sujetos como efecto de la eficacia del tratamiento. El principal inconveniente proviene de la secuencia en la presentación de las situaciones experimentales, y puede darse “aprendizaje” de una situación experimental antes de proceder a la siguiente.

Uno de los diseños intrasujeto más utilizados, sobre todo en la práctica clínica, es el denominado diseño DE CASO ÚNICO (N=1). No obstante, algunos autores incluyen este diseño dentro de la categoría de diseños cuasi-experimentales. El diseño de caso único no necesariamente implica que haya de aplicarse al estudio de un caso aislado, sino que el interés metodológico se focaliza en el análisis riguroso de los sujetos individuales por separado y no a partir de los indicadores estadísticos del grupo.

El interés por lo idiográfico, característico de la modificación de conducta, ha conducido a la elaboración de los diseños de caso único, ya que desde este enfoque se resalta la evaluación y tratamiento individualizado de cada paciente, contribuyendo, por lo tanto, a desarrollar este tipo de diseños (Echeburua, 1993). En la actualidad los diseños de caso único están ligados, sobre todo, al análisis conductual aplicado, tanto en el ámbito de las intervenciones terapéuticas como en el de la investigación básica.

Este tipo de diseño se utiliza preferentemente para evaluar los cambios de comportamiento, debidos a la influencia de un tratamiento, en un sujeto. Cobró auge a partir de los años setenta, al aumentar el interés por verificar la eficacia de los tratamientos en relación a la influencia sobre el sujeto de las variables específicas de la interacción del sujeto con el ambiente. Una de sus ventajas es que permite obtener feedback inmediato (retroalimentación) sobre los efectos del tratamiento, permitiendo la posibilidad de introducir variaciones a igual ritmo (Maganto, 1990, Ortiz, 1994).

Una característica fundamental de este método es que las medidas dependientes se evalúan repetidamente durante las fases del tratamiento. La repetición de las medidas va a aportar información respecto a la estabilidad o cambios tendenciales en la conducta de los sujetos, y permite sacar inferencias respecto a la línea de base.

Se considera de especial importancia en esta modalidad de diseño la definición precisa de las variables (dependientes, independientes, extrañas) y de las condiciones en las cuales

tiene lugar el experimento. Para determinar la validez interna de los resultados obtenidos se utiliza la replicación (repetición del experimento en condiciones lo más similares posibles), siendo los principales tipos de replicación la “intrasujeto” y la “intersujeto”. Los diseños de caso único permiten hacer frente a algunas de las limitaciones de los diseños intersujeto a nivel teórico, práctico y ético.

Sintetizando las características de los diseños de caso único, Ibañez (1993b) destaca las siguientes:

- Medición repetida y continuada de los sujetos en las variables dependientes.
- Registro de la conducta en la fase de línea base, es decir, antes de la intervención, con la finalidad de recoger información sobre la ocurrencia, intensidad, extensión y/o duración de la conducta.
- Intervención o tratamiento.
- Examen continuado de la variabilidad individual.
- Diseños de cambio rápido, dinámicos e interactivos, ya que el examen constante de los datos permite tomar decisiones rápidas sobre la conveniencia o no de modificar la intervención.
- Los datos están sometidos a una secuencia temporal e intrafase que se adecuan a modelos de series temporales ininterrumpidas.

Los diseños de caso único pueden agruparse en tres grandes bloques (Arnau, 1994, 1996b; Botella, 1986; Echeburua, 1993; González Almendros, 1987) que corresponden a las modalidades que más se utilizan en Psicología: Los “diseños intraseries”, “los diseños interseries”, y “las series combinadas”.

- Los “diseños intraseries” evalúan los cambios que aparecen en el sujeto en las distintas fases del tratamiento. Son los más utilizados en la modificación de conducta.
- Los “diseños interseries” tienen como objetivo comparar el efecto de dos o más tratamientos a través del tiempo. La modalidad de aplicación más común es la de tratamientos alternantes y simultáneos.
- La estrategia de las “series combinadas” de N=1, es una combinación de las dos anteriores. El investigador compara los resultados dentro de cada serie y entre ellas.

Como principales ventajas de los diseños de caso único se señalan su mayor validez interna, y su utilidad en la clínica y en el campo de la educación especial (Olivar, 2001). Entre sus limitaciones se encuentran las referidas a su validez externa en cuanto a la posibilidad de generalización de los resultados; la imposibilidad de investigar hipótesis etiológicas cuando los que se supone son factores causales no pueden ser manipulados experimentalmente; y limitaciones de tipo ético cuando se investigan las variables que causan un trastorno, ya que la única forma de demostrar científicamente las causas del mismo consistiría en provocarlo en sujetos sanos.

4.2.2.2. Diseño intergrupo

El diseño intergrupo, en su forma más simple, consiste en la aplicación a dos grupos de tratamientos diferentes (por manipulación de la variable independiente). A esta variable, la única que se diferencia entre los dos grupos, se atribuyen los efectos resultantes.

Los diseños experimentales de grupos se utilizan en modificación de conducta en la evaluación de la eficacia de las técnicas terapéuticas, en el análisis de los componentes activos de las mismas, así como en los estudios sobre la eficacia comparativa de distintos procedimientos terapéuticos. En todos estos casos los resultados se expresan a través de porcentajes de éxito (Echeburua, 1993; Labrador, 1982; Ruiz, 1987).

En estos diseños, los más utilizados son los diseños de grupo de control y los diseños factoriales, el efecto del tratamiento se evalúa por medio de la comparación entre grupos en lugar de hacerlo por medio de la evaluación de un sujeto a través del tiempo. Los sujetos se asignan aleatoriamente a los grupos experimentales y al grupo de control para reducir la probabilidad de que los grupos difieran antes del tratamiento. Antes de iniciar la intervención se establece la equivalencia entre los grupos, al menos respecto a la variable dependiente, para que ningún factor distorsionador previo contamine la interpretación de los resultados tras el tratamiento (Echeburua, 1993). El problema fundamental de esta modalidad de diseño radica en el control y condición de las variables.

Dentro de estos diseños intergrupo se puede diferenciar los diseños de grupo de control y los diseños factoriales. En los diseños de *grupo de control* se evalúa la conducta problemática antes de iniciarse el tratamiento y al finalizar el mismo, tanto en los grupos experimentales como en el de control. Posteriormente los grupos experimentales reciben los distintos tratamientos puestos a prueba, mientras que el grupo de control no recibe tratamiento alguno. La evaluación de la eficacia de un tratamiento depende de un cuidadoso control experimental sobre las hipótesis rivales que pueden dar cuenta del cambio producido, y debido a ello los grupos de control adquieren gran significación (Echeburua, 1993; Llavona, 1986).

En sus formas más complejas, en función del número de variables independientes que se quieran manipular, el diseño intergrupo recurre al diseño factorial. Los *diseños factoriales* intentan comparar los efectos simultáneos de dos o más variables independientes. Estos diseños son diseños de grupo de control en los que se examinan dos o más variables al mismo tiempo y se puede, por tanto, comprobar el efecto que sobre una(s) medida(s) dependiente(s) tiene la manipulación (por ejemplo, la relajación) o la introducción de dos o más variables independientes, así como el efecto interactivo entre las mismas (Ruiz, 1987). Estos diseños se hacen más complejos a medida que se incrementa el número de variables independientes que se quieran manipular (modalidades de tratamiento) y la diversidad de los efectos conductuales que conlleven (Campbell y Stanley, 1963/1988; Cook y Campbell, 1981).

Los diseños factoriales son útiles porque permiten averiguar la manera en que las variables independientes se combinan o actúan recíprocamente para influir sobre la conducta, y porque los efectos de diversas variables pueden examinarse al mismo tiempo en un solo experimento. En el campo de la modificación de conducta la principal dificultad radica en reunir las muestras de pacientes y disponer del suficiente número de terapeutas para que sea posible llevar a cabo los distintos tipos de intervenciones. Un estudio detallado de este tipo de diseños figura en trabajos como los de Cook y Campbell (1979) o Pereda (1987).

4.2.3. EL MÉTODO OBSERVACIONAL

La observación cuidadosa que describa con precisión acontecimientos de interés, ha sido el punto de partida de toda ciencia. La observación además de ser una actividad de la vida diaria, es el instrumento primordial de la investigación científica. La observación ha sido considerada por numerosos autores como la estrategia fundamental del método científico. En lo

que se refiere a la Evaluación Psicológica, cualquier procedimiento de evaluación se realiza a través de una u otra forma de observación (Anguera, 1981, 1994b; Brantley y Bruce, 1986; Fernández-Ballesteros, 1979, 1980, 1983, 1992c, 2004e).

Con la finalidad de delimitar la diferenciación entre método observacional y técnicas observacionales, Riba (1991/1993) destaca: “El *método observacional* puede verse como aquella competencia cognitiva, ligada a operaciones conductuales, que permite elaborar planes de investigación observacional, es decir, planes encuadrados en el esquema de rasgos observacionales: el experimento de campo informal, la observación de laboratorio y la observación de campo estricta. Una *técnica observacional*, en cambio, puede ser cualquiera de las que permiten llevar a la práctica sucesivas fases de ejecución de dichos planes, dentro de la gama de elecciones existentes en cada una de ellas”. Este investigador distingue varios tipos de técnicas:

- Técnicas de selección de unidades de observación y construcción de sistemas e instrumentos de observación, vertebrados sobre dichas unidades.
- Técnicas de selección de las muestras de sujetos.
- Técnicas de muestreo temporal o conductual.
- Técnicas de cuantificación y medición.
- Técnicas de control y calidad de datos (fiabilidad, validez, capacidad de generalización), tanto en la vertiente de la amortiguación de fuentes de error como en la de evaluación del mismo.
- Técnicas de agrupación y reducción de datos y análisis de los mismos.

La observación cuidadosa cuando se utiliza para describir de forma precisa acontecimientos de interés está considerada como una estrategia particular del método científico, de hecho se estima que ha sido el punto de partida de toda ciencia. La observación, además de ser una actividad de la vida diaria, es el instrumento primordial de la investigación científica. Según Anguera (1990ab), la observación es una estrategia del método científico cuyo objetivo es la cuantificación del comportamiento espontáneo tal y como éste se produce en situaciones no preparadas. Según Balluerka y Vergara (2002), el realismo es la condición esencial que caracteriza a la metodología observacional. El investigador cuando utiliza esta estrategia registra la información en su marco natural, observando sistemáticamente a un sujeto o a un grupo de sujetos, sin realizar ningún tipo de intervención sobre la situación observada. Hay que distinguir entre la observación como método y como técnica, estrategia de recogida de datos cuyo objetivo es el suministro de información (Delgado y Prieto, 1997).

No obstante, no toda observación se puede considerar científica. Para que la observación sea científica ha de ser sistemática, esto es, se ha de hacer de forma tal que pueda ser replicable por otro investigador. Entendida de tal forma, la observación consiste en una recolección de datos basada en el registro de comportamientos de una unidad de análisis (personas, conductas, sucesos...), y cuya finalidad es obtener una información comprensiva de los fenómenos que se producen en dichas unidades de análisis. Para ello se requiere especificar claramente sus objetivos y procedimientos dando respuesta a qué, cómo y con qué técnicas se llevarán a cabo. La idea esencial es que a través de la observación sistemática se puedan extraer regularidades de los datos, a fin de que permitan acceder a conclusiones interpretativas objetivas.

En la observación generalmente, es indispensable que el investigador fije previamente (Sarafino y Armstrong, 1988): a) las conductas específicas a registrar, b) cómo serán registradas,

es decir, cuáles serán los sistemas de codificación y las formas de recopilación de datos, c) establecimiento de recursos auxiliares, d) que la conducta puede ser interpretada fiable y consistentemente por varios observadores, y e) cómo serán transformados los datos en unidades cuantificables y cómo serán analizados. Por ello, Anguera (1996ab) define esta modalidad metodológica como el procedimiento encaminado a articular una percepción deliberada de la realidad, de forma que a través del registro objetivo, sistemático y específico de la conducta, se obtengan resultados significativos sobre el objeto de estudio. Servir a un objetivo ya formulado, ser planificada sistemáticamente, la optimización de los datos recogidos y la adecuación de la estrategia de análisis en función del objetivo son los requisitos que le dan entidad de método científico.

Para que se pueda llevar a cabo una recogida de datos de forma sistemática debe estar perfectamente delimitado: 1) *Qué observar* (niveles de análisis, marco teórico, definición de las categorías de observación, medidas: ocurrencia, frecuencia, latencia, duración e intensidad); 2) *Cómo observar* (observación natural, observación estructurada, experimento de campo, observación participante); 3) *Cuándo observar* (muestreo sistemático, muestreo asistemático, registro continuo, registro de intervalos); 4) *A quién observar* (delimitación de la muestra); y 5) *Dónde observar* (situaciones clínicas o naturales). Además, la observación como cualquier otro método científico, debe cumplir los requisitos de fiabilidad y validez.

En opinión de Anguera (1989) las fases del método observacional coinciden con las de la investigación científica en general, puesto que se trata de un proceso o secuencia de actividades encaminadas a ampliar el ámbito de nuestros conocimientos. Estas fases son: formulación de un problema, recogida de datos, análisis e interpretación de los datos observacionales, y comunicación de los resultados.

Los estudios observacionales suponen una recogida de datos basada en el registro de comportamientos de una unidad de análisis: personas, conductas, sucesos, etc. Su objetivo es obtener una información comprensiva de los fenómenos que se producen en dichas unidades de análisis. Requieren especificar claramente sus objetivos y procedimientos dando respuesta a qué, cómo y con qué técnicas se llevarán a cabo.

Los métodos observacionales son la selección, provocación, registro y codificación del conjunto de conductas y ambientes que implican a los organismos "in situ" y que concuerda con los objetivos empíricos. Esta definición pone de manifiesto los elementos fundamentales de casi todos los tipos de sistemas de observación:

- El observador debe seleccionar aquellas personas, categorías, sucesos o períodos de tiempo observables.
- Debe decidir si se provocarán o no las situaciones observables o si esperará que vengan de una manera natural.
- Establecerá un sistema de registro de observación.
- Establecerá un sistema de codificación de forma que las observaciones puedan operativizarse.

La observación puede ser utilizada tanto en diseños correlacionales como experimentales, en situaciones naturales o con máximo grado de control, pudiéndose distinguir diferentes tipos de estudios observacionales. En función del tipo de diseño que se utilice en la fase de recogida de datos se da lugar a distintas clases de observación (Anguera, 1996b). Según el grado de estructuración la observación puede ser *sistemática* o *asistemática*, según el grado de

inferencia puede ser *directa* o *indirecta*, según el papel del observador se distingue entre observación participante y no participante, y según el lugar de observación hablamos de observación *natural* o *de laboratorio*.

La observación natural

La observación natural puede definirse como aquella que se realiza en el ambiente ordinario en el que se desenvuelve el sujeto sin que se produzca ningún tipo de mediación del evaluador en provocar las actividades objeto de estudio (Fernández-Ballesteros, 1992c). La observación naturalista se lleva a cabo en muy distintos ambientes, en el ambiente familiar, en el contexto escolar y en ambientes lúdicos y clínicos. A lo largo de la historia los psicólogos interesados en el aspecto evolutivo-genético han elaborado sistemas de observación que se dirigen hacia ciertas categorías específicas de la conducta que se considera que indican determinadas etapas o niveles de funcionamiento físico, cognitivo y social (Berstein y Nietzel, 1982). En el área de la psicología social ha sido un método que ha posibilitado el análisis de las acciones recíprocas que se desarrollan en los grupos. En esta línea surgió el sistema de codificación de las interacciones de Bales (1950). Así mismo, el deseo de observar los comportamientos en diversos contextos específicos (hospitales, cárceles, psiquiátricos, escuelas...) ha enriquecido el desarrollo del método y ha concluido con sistemas categoriales.

Como señala Maganto (1995b) las denominadas biografías de bebés son el exponente más extremo de observación naturalista, pero generalmente presentan múltiples limitaciones porque carecen de una observación sistemática de conductas, los observadores son los propios padres, no existen objetivos determinados de observación ni técnicas adecuadas de recogida de datos. Como contrapartida han aportado datos de gran relevancia e interés para investigaciones posteriores y han contribuido a desarrollar el método observacional. Producto de esta técnica es el método "estudio de casos" (Sarafino y Armstrong, 1988) calificado como biografía sistemática con base en los registros del historial del niño, siendo un método muy común en sujetos infantiles con problemas emocionales, psiquiátricos, fracaso escolar... También estos trabajos se han enriquecido progresivamente a través de sistemas de categorías de mayor precisión y operatividad.

La observación sistemática

La observación sistemática se asemeja al método científico acercándose al modelo de experimentación. Es más, la observación puede formar parte o estar inserta en un contexto experimental, de máxima artificialidad. Sin embargo, no debe confundirse con ella. Para Anguera (1985), una de las máximas promotoras del método, la observación y la experimentación son dos métodos que difieren entre sí, ya que resalta que en la primera el investigador considera los fenómenos tal como se presentan, sin modificarlos ni actuar sobre ellos, mientras que la experimentación implica una variación o perturbación provocada intencionalmente por el experimentador sobre las condiciones en las que se desarrolla el fenómeno.

Anguera (1990) entiende por observación sistemática el procedimiento encaminado a la percepción deliberada de una realidad conductual de forma que mediante su registro, codificación y análisis nos proporcione resultados significativos del sujeto en evaluación. Para Bakeman y Gotman (1986/1989) la observación sistemática es una vía específica de acceso a

la cuantificación de la conducta. Esta aproximación se refiere típicamente a la ocurrencia espontánea de la conducta que se observa en contextos naturales. El propósito consiste en definir de antemano varias modalidades de conducta (códigos conductuales) y solicitar entonces a los observadores que registren, cada vez que ocurra la conducta, el correspondiente código predefinido. El mayor interés se centra en entrenar a los investigadores de forma similar, dado que han observado la misma corriente de conducta.

Así, los pilares de la observación sistemática serían la utilización de catálogos predeterminados de códigos de conducta y observadores que demuestren una alta concordancia. Para estos autores, la observación sistematizada se basa en la precisión y exactitud, en la definición del sistema de categorías de observación realizada con dos criterios básicos: exhaustividad (se refiere a que cualquier comportamiento del ámbito considerado como objeto de estudio puede asignarse a una de las categorías) y exclusividad de un evento observacional (significa el no solapamiento de las categorías que componen un sistema, por lo que a cada comportamiento se le asignaría una y sólo una categoría).

La observación sistemática ha cobrado gran relevancia en el campo de la evaluación psicológica, presentando desarrollos importantes en cuanto a los objetivos y decisiones básicas a tomar en el método observacional (Riba, 1991/1993), al proceso de categorización (Anguera, 1991/1993; 1994b), a los sistemas de codificación (Bakeman y Gotman, 1986; Blanco y Anguera, 1991/1993), a la métrica y muestreo del registro observacional (Carreras, 1991/1993; Quera, 1991/1993), así como a la observación y análisis de la producción verbal de la conducta a través de documentos personales como autobiografías, diarios, cartas... (Behar, 1991/1993). Muchos investigadores están desarrollando estudios utilizando la metodología observacional tanto cualitativa como cuantitativa. Algunos se han centrado en las posibilidades de la metodología observacional en el estudio de conductas en el aula, empleando observación sistemática y análisis secuencial de los datos (Herrero, 1992), o en el análisis de las interacciones madre-hijo mostrándose una metodología adecuada para estos objetivos (Corboz-Warnery y Fivaz-Depeursinge, 1993). Otros estudios se han dirigido al análisis de la importancia de la calidad de los registros observacionales destacándose que la evaluación y la posible optimización de los registros es una etapa crucial en el proceso de la investigación educativa, ya condicionarán las conclusiones que se obtengan (Tojar, 1994).

La observación externa y la autoobservación

En función del grado de participación del observador en la investigación se pueden delimitar distintos tipos de observación. En general, la observación participante es un método de sondeo preliminar que permite elaborar hipótesis para ser contrastadas o verificadas posteriormente. Se utiliza preferentemente para observar las respuestas de los sujetos en actividades y situaciones de la vida cotidiana, para observar la interacción social en el seno de grupos e instituciones en donde el observador desempeña un papel activo. Este tipo de observación para Anguera (1985) tiene tres ventajas: a) mayor sensibilidad a la situación-escenario e interrelación en el grupo; b) mayor número de oportunidades de observación; y c) mayor posibilidad de acceso a datos ocultos a las personas ajenas al grupo. Como inconveniente se destaca el peligro de la subjetividad, la escasa estandarización, las dificultades de realizar réplicas y, por lo tanto, dificultades para la generalización de los resultados. Anguera (1994b) ha planteado una taxonomía en función del grado de participación

del observador, proponiendo 4 modalidades graduadas en base al continuum de la relación del observador y el observado:

- *Observación no participante:* Entre observador y observado no existe ningún tipo de relación.
- *Observación participante:* Existe una iniciativa del observador acerca del observado, como en un diálogo o en un entrevista clínica.
- *Participación-Observación:* El observador y el observado pertenecen a un grupo natural de sujetos, como diada madre-hijo, profesor-alumno.
- *Autoobservación:* En este continuum va reduciéndose la distancia entre observador y observado hasta llegar a la autoobservación, en que coinciden en una misma persona. Hay que distinguir entre autoobservación de conductas hetero-observables (número de cigarrillos fumados en un período de tiempo) y la autoobservación introspectiva referida a vivencias experimentadas por el observado, presentado esta última importantes riesgos de inferencia desmesurada y distorsión.

La observación experimentó un importante auge a partir de los años cincuenta, gracias a la modificación de conducta y la evaluación conductual. La observación conductual, sistemática, constituye el procedimiento básico de recogida de información al centrarse en conductas externas, planteándose como objetivo determinar y cuantificar la conducta o conductas problema, así como las variables mantenedoras y consecuentes (De la Puente, Labrador y García, 1995). Anguera (1999) expone el amplio campo de aplicación de la metodología observacional en cinco monografías, en las que se recogen diferentes investigaciones que han utilizado esta metodología en diferentes contextos y situaciones.

Anguera (1990a, 1996b, 2003ab) destaca como ventajas de este método: la capacidad para generar hipótesis en base a la observación de regularidades en la conducta del niño; la no subordinación a la necesidad de control, esto permite obtener mucha información de la conducta espontánea del niño en interacción con las variables de la situación en el contexto que se producen; juntamente con que se puede aplicar en casos donde no es posible utilizar otras alternativas metodológicas. Como limitaciones más importantes se destacan: la inconsistencia de los datos cuando se obtienen de manera puntual debido a posibles fluctuaciones, la dificultad de codificación de los acontecimientos cuando las situaciones son complejas; el riesgo de realizar inferencias en la interpretación de los datos, el problema de la fiabilidad interjueces y la validez de las observaciones.

4.3. ALTERNATIVAS METODOLÓGICAS POLÉMICAS

Derivadas de los distintos modelos y métodos utilizados en Evaluación Psicológica se han generado una serie de alternativas polémicas que corresponden con las dicotomías clásicas que tiene planteadas la Psicología Científica. Suelen numerarse con frecuencia cuatro controversias: Generalidad-Especificidad de la conducta, Idiográfico-Nomotético, Cuantitativo-Cualitativo, y Explicación-Comprensión (Fernández-Ballesteros, 1992a, 2004a; Garaigordobil, 1998).

4.3.1. GENERALIDAD - ESPECIFICIDAD DE LA CONDUCTA

Esta dicotomía se inscribiría dentro de las características diferenciales dadas por Barrios y Hartman (1986) como propias de la evaluación tradicional frente a la conductual. Según estos autores la denominada Evaluación Tradicional pretende predecir o evaluar la conducta en base a estructuras internas que se han ido constituyendo a lo largo de la vida del sujeto, dando estas estructuras estabilidad al comportamiento de esa persona, tanto a través del tiempo como en diversas situaciones. Por el contrario, desde la Evaluación Conductual más radical se pretende explicar la conducta en base a estímulos externos, y por lo tanto, se sostiene que el comportamiento es específico, y depende de determinados factores ambientales que concurren en el momento en que se produce esa conducta.

En relación a esta dicotomía, Westmeyer (1992) criticó estas consideraciones planteando dos argumentos:

- 1) La afirmación de que en el enfoque tradicional las causas de la conducta son intrapsíquicas no parece un retrato fiel de las posiciones psicométricas. Las típicas aproximaciones psicométricas son aproximaciones cuantitativas, basándose no en un modelo que afirma que la conducta está en función de rasgos internos de personalidad, sino que afirma que la conducta está en función de la interacción de la persona y la situación. A pesar de que tanto persona como situación son considerados de forma diferente desde distintos enfoques psicométricos, la interacción entre estos dos conceptos es común a todos ellos.
- 2) Con respecto al tema de la consistencia de la conducta Westmeyer aduce que si describimos de una forma muy estricta el componente situacional en un diagnóstico de tipo conductista, no podemos inferir nada de tal ecuación conductual, puesto que nunca más volverá a darse dicha situación. Si elegimos en cambio un tipo de descripción más abstracto y generalizado, deberíamos aceptar dentro de la ecuación conductista una cierta consistencia de la conducta. Además, en la concepción psicométrica de Catell, por ejemplo, no se espera que la conducta sea consistente a través del tiempo y de los "settings" (escenarios).

En consecuencia Westmeyer (1992) llega a la conclusión de que en los dos acercamientos a la evaluación la consistencia de la conducta es un tema empírico, y depende de varios factores cuyo impacto no puede ser juzgado a priori. En esta misma línea se sitúan los enfoques interaccionistas desde los cuales se han formulado modelos interactivos en base a los cuales, en opinión de Fernández-Ballesteros, (1992a) puede concluirse lo siguiente:

- La conducta de un sujeto en un momento dado está en función de las interacciones recíprocas entre las variables situacionales y personales (Bandura, 1987, Staats, 1980).
- Las variables personales u organísmicas (entre las cuales Fernández-Ballesteros, 1992, incluye el modo en el cual el sujeto percibe la situación en la que se encuentra) son a su vez dependientes de las interacciones y/o del proceso dialéctico habido entre el organismo y el medio ambiente durante la vida del sujeto y, por lo tanto, dependen de su historia del aprendizaje (Staats, 1975).
- La mayor especificidad o generalidad de la conducta de un determinado sujeto estará en función del balance entre los procesos de generalización y de discriminación habidos

durante su historia de aprendizaje. Esto implica que algunos repertorios de conducta son más estables que otros en ciertas ocasiones y para algunos sujetos (Magnusson, 1990; Mischel, 1990).

La conclusión más evidente está en la línea de preconizar un sujeto activo y actuante en un ambiente activo. Todo ello implica que habremos de evaluar variables ambientales, personales, además de la propia conducta objeto de estudio, y sus relaciones mutuas.

4.3.2. IDIOGRÁFICO-NOMOTÉTICO

En el fondo del debate subyace la asunción de que la ciencia no puede basarse en el caso único. A este respecto se invocan como paradigmáticas las palabras de Eysenck: “a la ciencia no le interesa el hecho único; lo único pertenece a la historia y no a la ciencia” (Eysenck, 1951).

El enfoque idiográfico se dedica al estudio de los fenómenos individuales, mientras que el nomotético se centra en el hallazgo de principios generales aplicables a los fenómenos que estudian. Tal división, así como el encuadre de la Psicología entre las ciencias idiográficas, ha provocado múltiples discusiones. La aproximación idiográfica ha sido elegida por los modelos clínicos y fenomenológicos, sin embargo el modelo psicométrico ha optado por una aproximación nomotética. En cuanto al modelo conductual es esencialmente idiográfico.

Para Fernández-Ballesteros, la Psicología Científica es, inequívocamente, una disciplina cuyo objetivo es la búsqueda de los principios generales aplicables a la conducta humana y, por tanto, es una ciencia claramente nomotética. Y ya que la Evaluación Psicológica ha sido definida como una subdisciplina de la Psicología, por fuerza, ha de basarse en los principios de la ciencia de la que surge y a la que sirve. Es más, su objetivo consiste, precisamente, en el análisis de cómo se cumplen y se organizan los principios psicológicos básicos en un sujeto específico. Ahora bien, esto quiere decir que su objeto fundamental reside en el caso concreto o en un ente individual (Fernández-Ballesteros, 1992a, 2004a).

Como se deja entrever Fernández-Ballesteros (1980, 1992a, 2004a) adopta una posición sintetizadora, subrayando: 1) Los objetivos de la evaluación son, fundamentalmente idiográficos, en el sentido de que éstos se centran en el estudio científico del comportamiento de un sujeto; 2) Sin embargo, nuestra disciplina se basa en los hallazgos de una disciplina nomotética y, por lo tanto, en los resultados establecidos para los distintos hechos psicológicos; y 3) Los métodos utilizados en Evaluación son los mismos de la Psicología nomotética y encuadrables dentro de una perspectiva científica.

Así mismo, Forns (1985, 1993) señala que “la evaluación se da en una encrucijada en la que confluyen la información múltiple procedente de los dominios del conocimiento (conocimiento nomotético) y la proveniente del caso concreto o situación específica. La toma de decisión diagnóstica es posible cuando se da cierta concordancia y coherencia explicativa (de matiz causal o funcional) entre los datos procedentes de ambas fuentes (lo teórico y lo idiográfico). La evaluación apela a ambos tipos de conocimientos, lo nomotético y lo idiográfico, y el acto clínico consiste en el adecuado engranaje entre ambos (Forns, 1993).

En esta dirección integradora se puede afirmar que toda evaluación implica inevitablemente a ambos enfoques: por un lado, toda evaluación, por cuanto trata de diferenciar entre eventos, es inherentemente idiográfica; por otro lado, el proceso de diferenciación

inevitablemente implica el uso de elementos nomotéticos tales como las abstracciones, clasificaciones, generalizaciones y teorías. La cuestión crítica, por consiguiente, no es si es mejor la evaluación nomotética o idiográfica. La cuestión clave es saber qué unidades de análisis nomotéticas pueden resultar más útiles para describir, predecir y explicar el particular evento único que nos interesa.

Así, es posible estar orientado hacia los rasgos y a la vez idiográficamente y a la inversa. La evaluación conductual idiográficamente pura, en palabras de Cone (1986), es un capricho científico más que una realidad, especialmente cuando se pretende asimilar lo idiográfico con procedimientos inductivos y nomotético con procedimientos deductivos. En la actualidad es difícil mantener la dicotomía, ya que la evaluación puramente idiográfica implicaría descontextualizar completamente al sujeto de evaluación. Implícita o explícitamente, toda evaluación conlleva una idea de comparación con el grupo social de pertenencia.

4.3.3. CUANTITATIVO-CUALITATIVO

Otra dicotomía que ha sido planteada en el contexto de la evaluación psicológica es la de lo cuantitativo versus lo cualitativo. Desde el enfoque cualitativo, denominado también clínico, se pretende realizar un análisis global del sujeto a examen, enfatizando la libertad de juicio del profesional y dejándole, por tanto, amplios márgenes tanto para la elección de las situaciones de examen como para el tratamiento de los datos procedentes de ellas. Sin embargo, desde el enfoque cuantitativo el objetivo sería una mayor precisión a la hora de evaluar las respuestas del sujeto, unas mayores exigencias de estandarización de las situaciones de prueba y una ponderación y elaboración mecánica de los datos, así como procedimientos uniformes para la comunicación de los resultados. Así, la perspectiva cualitativa defiende el estudio global y cualitativo del sujeto, siendo factores determinantes la capacidad intuitiva del clínico, su experiencia y empatía; sin embargo, la alternativa cuantitativa defiende el rigor numérico, la cuantificación, la fórmula matemática aplicable a una determinada manifestación conductual y la replicabilidad del proceso por cualquier juez.

A partir, sobre todo, de las investigaciones de Cronbach (1960/1972), la polémica ha ido disminuyendo y se produce una tendencia a posiciones integradoras. El debate se ha ido desradicalizando y se tiende a la integración de posturas. Actualmente podemos considerar que ambos enfoques son útiles porque: 1) El subjetivismo inherente al enfoque cualitativo ha sido superado por la estadística no paramétrica que permite contrastar empíricamente datos cualitativos; 2) Las investigaciones realizadas en la perspectiva del enfoque cualitativo tienden a abandonar inferencias y ambigüedades descriptivas; y 3) Se ha puesto de manifiesto la posibilidad de cuantificación de los resultados obtenidos a través de la evaluación cualitativa.

En la actualidad se acepta que el diagnóstico combina datos cuantitativos y cualitativos, por lo que no cabe hablar de dilema, y ambos enfoques cuentan con aspectos positivos. Por ello, y en la misma dirección que la indicada por Fernández-Ballesteros (1992a, 2004a), considero que esta polémica debe ser abandonada por los siguientes argumentos:

- 1) Lo artificial de la misma, por cuanto la mayor parte de los psicólogos que realizan una evaluación utilizan datos cuantitativos y cualitativos. Existen múltiples posibilidades de combinación de datos en las diferentes fases de la evaluación: el evento a evaluar, el tipo de procedimiento de recogida de información, la elaboración de ésta, y la inferencia que el

psicólogo realice sobre aquél. Así, el tipo de evento a examen puede ser máximamente cualitativo, subjetivo e interno (una atribución, un sentimiento o imagen...) u objetivo o externo (un evento manifiesto de cualquier tipo). También el procedimiento de recogida de información puede ser altamente cualitativo (un diario u otro documento personal) o cuantitativo (una puntuación numérica procedente de un registro mecánico...). Por otra parte, el tipo de manipulación a la que el evaluador somete los datos recogidos puede ser, fundamentalmente cualitativa (como ocurre en un análisis de contenido) o cuantitativa (cuando los datos se procesan estadísticamente). Finalmente, de los resultados obtenidos el evaluador puede realizar combinaciones cualitativas de un alto nivel de abstracción o bien puede mantenerse con estricta referencia a los datos cuantitativos. Las combinaciones de estas cuatro fuentes de variación, con sus distintos grados, pueden producir innumerables alternativas, lo cual diluye la polaridad de la doble alternativa.

- 2) El abandono en la práctica del subjetivismo a lo largo de todo el proceso de evaluación, incluso por los propios psicólogos partidarios del análisis clínico. En este momento es mantenido tan sólo por los enfoques fenomenológico-humanistas.
- 3) Los avances metodológicos realizados en torno al estudio del caso único permiten el mayor rigor y control experimental junto con la utilización de datos cualitativos. Así mismo, los diversos sistemas expertos desarrollados para el diagnóstico psicológico demuestran la posibilidad de introducir datos cualitativos en heurísticos computacionales.

Aunque desde esta perspectiva se plantea la posibilidad de combinar lo cuantitativo con lo cualitativo esto no implica aceptar un alto grado de subjetivismo, ya que la Evaluación Psicológica debe permitir la posibilidad de que otros evaluadores repliquen lo realizado. Como ya señaló Cronbach (1985) es importante que los psicólogos inventen métodos para un procesamiento riguroso de la información cualitativa y que de igual manera dependan cada vez más de lo cualitativo para dar sentido a lo cuantitativo. Por ello, lo cuantitativo sin lo cualitativo no tiene razón de ser.

A partir de aquí, es obvio, como afirmaba Thorndike (1918), que todo cuanto existe, existe en alguna medida, por tanto, es susceptible de cuantificación. Pero, así mismo, todo cuanto existe, existe de determinada manera, y por lo tanto, es susceptible de cualificación. El problema real se centra, más que en la elección forzosa y excluyente de alguna de estas dos dimensiones definitorias del fenómeno, se centra en la correcta operativización de estas dos dimensiones definitorias del objeto o fenómeno. En concreto, la cuestión principal se centra en decidir, no si hemos de cuantificar o cualificar el fenómeno, sino en determinar *qué tipo de parámetro* debe emplearse para optimizar la validez predictiva de nuestra evaluación o su poder explicativo. Si nos centramos a modo de ejemplo en la inteligencia, ¿qué unidad de análisis será la más adecuada en base a unos determinados objetivos?: ¿el número de problemas que es capaz de resolver el sujeto, el tiempo que tarda en resolverlos, la cantidad de tensión psicofísica que sigue el sujeto para resolverlos, o, acaso, sus estrategias resolutorias....?.

4.3.4. EXPLICACIÓN-COMPREENSIÓN

Una derivación de la polémica sobre lo idiográfico versus lo nomotético, anteriormente planteada, fue defendida por Dilthey, quien sostuvo que la explicación era el objetivo de las

ciencias naturales, mientras que la comprensión lo era de la Psicología, considerada ésta como “una ciencia del estudio del espíritu”. La comprensión desde la perspectiva adoptada por Dilthey, implica la búsqueda del sentido de la acción humana en contraposición con la investigación causal de los fenómenos psicológicos. Así, contrapone el estudio científico de la conducta humana versus la comprensión empática de los sujetos evaluados.

En la actualidad, y en la misma dirección que sucedía con las polémicas anteriormente planteadas, tendencialmente los autores adoptan posiciones integradoras, rechazando las posturas extremas planteadas en torno a esta polémica. Por una parte, la que representa una ciencia dura, desinteresada de la comprensión y guiada por el deseo de verificar rigurosamente las predicciones deducidas de la teoría, una interpretación, a su juicio errónea, de lo que ocurría en la Física y la Química, y adoptada con entusiasmo por muchos psicólogos. Por otra, la seguida por un pequeño número de psicólogos, adheridos a unos argumentos filosóficos que les permitían rechazar el método científico y el control para, de esta manera, dar rienda suelta a la especulación irresponsable y a una intuición sin disciplina. Este tipo de comprensión no explicativa es más propia del artista que de un científico.

En síntesis: Desde este proyecto se entiende la Evaluación Psicológica como una subdisciplina de la Psicología Científica, y por lo tanto se debe considerar incluida en una perspectiva científica. Su finalidad radica en comprobar si las leyes generales establecidas por la Psicología se verifican en personas concretas. De este modo encontramos que posee unos objetivos de carácter idiográfico dentro de una ciencia nomotética. Partiendo de estos supuestos, estaría de acuerdo con los autores que adoptan posiciones integradoras en relación a las cuatro polémicas.

4.4. RESUMEN: PRINCIPALES MÉTODOS Y ALTERNATIVAS POLÉMICAS

Con la finalidad de clarificar algunas de las cuestiones abordadas en este capítulo, a continuación se sintetizan las características de los principales métodos utilizados en evaluación psicológica (Cuadro 9), así como diversas alternativas metodológicas polémicas frente a las que se presenta una integración personal (Cuadro 10).

Cuadro 9. Principales métodos en Evaluación Psicológica

PRINCIPALES MÉTODOS EN EVALUACIÓN PSICOLÓGICA	
Definición	Método: Cualquier procedimiento sistemático que tenga por finalidad llegar a un objetivo, y tanto el fin como la forma deben estar establecidos de antemano y pueden ser replicados por cualquiera. Principales métodos: Correlacional, Experimental y Observacional
Método Correlacional	• Correlación: medición estadística de la relación recíproca entre variables que indica dirección y magnitud de dicha relación. El grado de correlación posibilita la predicción de los resultados que se pueden encontrar en una variable (criterio) a partir de los resultados obtenidos en otra (predictor). El predictor no se puede tomar como variable causal. • Fija su punto de interés en las diferencias entre individuos, y la clasificación de los sujetos en función de estas diferencias. • Busca establecer relaciones de concomitancia entre los hechos tal y como se producen naturalmente, sin la manipulación del experimentador. • Método dirigido primordialmente al estudio de la estructura de la inteligencia y la personalidad. • Ha conducido a establecer conceptos claves en la investigación estadística tales como fiabilidad, validez o varianza de error.
Método Experimental	• Estudia los fenómenos bajo condiciones de estricto control. • No tienen interés en las diferencias individuales. • Busca establecer relaciones de causalidad que le permita imputar a la manipulación de las variables independientes los cambios observados en las variables dependientes. El experimentador elabora una hipótesis en la que presupone que X es la causa de Y. Para su comprobación actúa sistemáticamente sobre la variable X con el propósito de analizar su efecto en la variable Y. La manipulación de la variable que se hipotetiza como causa se denomina variable independiente (VI). La conducta o cambios en la conducta que se observan como efecto de lo anterior, se denomina variable dependiente (VD). • La metodología experimental persigue como objetivo estimar la magnitud del efecto de las variables de tratamiento o variables independientes o causales y se fundamenta en tres pilares: aleatorización, control e hipótesis de causalidad. El control, la manipulación experimental y la aleatorización en la selección de la muestra y asignación de los sujetos a grupos son las características esenciales de este método • Clasificación de los diseños experimentales tomando como referencia dos criterios: el número de sujetos y el número de variables. Según los Sujetos: • Un sólo grupo de sujetos: Se efectúan comparaciones entre dos o más tratamientos realizados a un mismo grupo de sujetos aunque desde el punto de vista de la lógica experimental, es como si tuviéramos dos o más grupos distintos. Este tipo de diseño se denomina INTRASUJETOS. Un caso particular de estos, muy discutido desde el punto de vista metodológico, es cuando sólo se utiliza un único sujeto. A este caso se le denomina diseño DE CASO ÚNICO (N = 1). • Más de un grupo de sujetos, denominándose diseños INTERSUJETOS. Estos incluyen: Diseños de dos grupos aleatorios, Diseños de dos grupos equiparados, y Diseños de más de dos grupos.

	Según las variables: • Una sola VD, es decir, un sólo efecto, que se denomina diseño UNIVARIADO. Cuando en este diseño se estudia una sola causa, es decir, una sola variable independiente se denomina diseño de COMPARACIÓN, mientras que cuando existen más de una variable independiente se denomina diseño FACTORIAL. • Más de una VD, es decir, más de un efecto, que se denomina diseño MULTIVARIADO.
Método Observacional	Metodología observacional y técnicas observacionales • El <i>método observacional</i> puede verse como aquella competencia cognitiva, ligada a operaciones conductuales, que permite elaborar planes de investigación observacional, es decir, planes encuadrados en el esquema de rasgos observacionales: el experimento de campo informal, la observación de laboratorio y la observación de campo estricta. Los métodos observacionales son la selección, provocación, registro y codificación del conjunto de conductas y ambientes que implican a los organismos "in situ" y que concuerda con los objetivos empíricos. • Una <i>técnica observacional</i>, en cambio, puede ser cualquiera de las que permiten llevar a la práctica sucesivas fases de ejecución de dichos planes, dentro de la gama de elecciones existentes en cada una de ellas. Observación y Evaluación Psicológica La observación cuidadosa cuando se utiliza para describir de forma precisa acontecimientos de interés está considerada como una estrategia particular del método científico, de hecho se estima que ha sido el punto de partida de toda ciencia. En lo que se refiere a la Evaluación Psicológica, cualquier procedimiento de evaluación se realiza a través de una u otra forma de observación Tipos de observación La observación puede ser utilizada tanto en diseños correlacionales como experimentales, en situaciones naturales o con máximo grado de control, pudiéndose distinguir diferentes tipos de estudios observacionales. • <i>Observación naturalista</i>: Ambiente natural sin mediación del evaluador • <i>Observación sistemática</i>: Se acerca al modelo de experimentación, puede formar parte de ella, pero no deben confundirse ya que en la observación el investigador considera los fenómenos tal como se presentan, sin modificarlos ni actuar sobre ellos. Procedimiento encaminado a la percepción deliberada de la realidad conductual de forma que mediante su registro, codificación y análisis nos proporcione resultados significativos del sujeto de evaluación. Vía de acceso específica de cuantificación de la conducta. • <i>Observación externa - Autoobservación</i>: 4 modalidades graduadas en el continuum de la relación observador-observado: observación no participante, observación participante, participación-observación, autoobservación. Parámetros de una observación Para que se pueda llevar a cabo una recogida de datos de forma sistemática debe estar perfectamente delimitado: 1) <i>Qué observar</i> (niveles de análisis, marco teórico, definición de las categorías de observación, medidas: ocurrencia, frecuencia, latencia, duración e intensidad); 2)

	<i>Cómo observar</i> (observación natural, observación estructurada, experimento de campo, observación participante); 3) <i>Cuándo observar</i> (muestreo sistemático, muestreo asistemático, registro continuo, registro de intervalos); 4) <i>A quién observar</i> (delimitación de la muestra); y 5) <i>Dónde observar</i> (situaciones clínicas o naturales). Además, la observación como cualquier otro método científico, debe cumplir los requisitos de fiabilidad y validez.

Cuadro 10. Alternativas metodológicas polémicas en Evaluación Psicológica

ALTERNATIVAS METODOLÓGICAS POLÉMICAS	
Generalidad Especificidad De la conducta	• <i>Evaluación Tradicional</i>: Explica la conducta en base a estructuras internas que se han ido constituyendo a lo largo de la vida del sujeto, dando estas estructuras estabilidad al comportamiento tanto a través del tiempo como en diversas situaciones. • <i>Evaluación Conductual</i>: Explica la conducta en base a estímulos externos, y por lo tanto, el comportamiento es específico y depende de determinados factores ambientales que se dan en ese momento. • <i>Posición integradora</i>: 1) la conducta en un momento dado es función de las interacciones recíprocas entre las variables situacionales y personales; 2) las variables personales y orgánicas son a su vez dependientes del proceso dialéctico habido entre el organismo y el medio ambiente durante la vida del sujeto; y 3) algunos repertorios de conducta son más estables que otros en ciertas ocasiones y para algunos sujetos.
Idiográfico Nomotético	• <i>Enfoque nomotético</i>: Se centra en el hallazgo de principios generales aplicables a los fenómenos que se estudian. Elegido por el modelo psicométrico. • <i>Enfoque idiográfico</i>: Se centra en el estudio de los fenómenos individuales. Elegido por los modelos clínicos, fenomenológicos, conductuales... • <i>Posición integradora</i>: El objetivo de la evaluación psicológica consiste en el análisis de cómo se cumplen y organizan los principios psicológicos básicos en un sujeto específico. Por lo tanto, toda evaluación implica a ambos enfoques, apela a ambos tipos de conocimientos, y el acto clínico consiste en el adecuado engranaje de ambos.
Cuantitativo Cualitativo	• <i>Enfoque cualitativo</i>: Denominado también clínico. Realiza un análisis global del sujeto a examen. Deja amplios márgenes a la libertad de juicio del profesional, a su capacidad intuitiva, experiencia y empatía. • <i>Enfoque cuantitativo</i>: Denominado también actuarial. Implica mayor precisión a la hora de evaluar las respuestas del sujeto, mayores exigencias de estandarización de la situación de prueba, ponderación y elaboración mecánica de los datos, y procedimientos uniformes para la comunicación de los resultados. Defiende el rigor numérico, la cuantificación, la fórmula matemática aplicable a una determinada manifestación conductual y replicabilidad del proceso. • <i>Posición integradora</i>: La evaluación psicológica puede combinar el método cuantitativo y cualitativo. En la actualidad: (1) ha sido superado el subjetivismo del enfoque cualitativo gracias a la estadística no paramétrica que permite contrastar empíricamente datos cualitativos; (2) en el enfoque cualitativo se evitan las inferencias y ambigüedades descriptivas; y (3) se pueden cuantificar los resultados obtenidos a través de la evaluación cualitativa.

Explicación Comprensión	• <i>Explicación</i>: Búsqueda de las causas de los fenómenos psicológicos. • <i>Comprensión</i>: Búsqueda del sentido de la acción humana. • <i>Posición integradora</i>: La evaluación psicológica debe realizar una comprensión explicativa.

CAPÍTULO 5. EL PROCESO DE LA EVALUACIÓN PSICOLÓGICA

- 5.1. Introducción
- 5.2. El proceso en función de las metas de evaluación
- 5.3. Fases y tareas del proceso de evaluación psicológica
 - 5.3.1. La perspectiva correlacional: Proceso descriptivo-predictivo
 - 5.3.2. La perspectiva experimental: Proceso interventivo-valorativo
 - 5.3.3. Reflexiones sobre la evaluación psicológica aplicada como proceso correlacional versus experimental
- 5.4. Guías para el proceso de evaluación (GAP)
- 5.5. Condicionantes que intervienen en el proceso de la evaluación psicológica
 - 5.5.1. Variables del evaluador
 - 5.5.2. Variables del evaluado
 - 5.5.3. Variables de la situación
 - 5.5.4. La calidad del proceso de evaluación psicológica
 - 5.5.4.1. La eficacia de los evaluadores en tareas de clasificación, predicción, prescripción de tratamiento...
 - 5.5.4.2. La relevancia de la capacidad de metacognición del evaluador
- 5.6. Resumen: Fases, objetivos, tareas y técnicas en el proceso de la evaluación psicológica

5.1. INTRODUCCIÓN

El profesional de la evaluación psicológica, con independencia de su adscripción teórica a un modelo concreto, cuando se enfrenta a la práctica diaria precisa de unas pautas de actuación que le permitan organizar su labor de una manera sistemática para conseguir unas metas satisfactorias. Hablamos en definitiva, de la existencia del proceso como un intento de organizar los diversos pasos que se suceden a lo largo de la praxis evaluadora.

Sin embargo, no existe un proceso único. Esta estructuración organizativa, o sea estas fases, en parte van a estar mediadas por algunos factores, como por ejemplo, el marco teórico de referencia del evaluador, ya que este marco condiciona su forma de ver y entender al individuo objeto de estudio y los instrumentos de evaluación a utilizar para llevar a cabo la evaluación. De esto se deduce la existencia de diversas pautas organizativas, diversas opciones, como lo han subrayado diversos estudios.

Actualmente se reconoce que el diagnóstico es un proceso complejo y dinámico, que ocurre a lo largo del tiempo y que implica una serie de fases con unos objetivos y procedimientos determinados que le llevan al fin deseado. Pero no hay una unificación de criterios respecto a cómo debería desarrollarse, existiendo diferentes alternativas en función tanto de las metas para las que se realiza (diagnóstico, selección, orientación, tratamiento) como de los sujetos implicados (una persona, un grupo, una institución) y las diferentes conceptualizaciones del proceso (por ejemplo, el proceso como procedimiento correlacional o como procedimiento experimental), estableciéndose un número variable de fases y procedimientos.

Por lo tanto, nos hallaríamos ante una multiplicidad de procesos de evaluación, y precisamente ante esta disyuntiva planteada, se desarrollará el tema planteando las fases desde la perspectiva de la evaluación tradicional (metodología correlacional), desde el enfoque conductual (metodología experimental).

5.2. EL PROCESO EN FUNCIÓN DE LAS METAS DE EVALUACIÓN

La amplitud y diversidad de objetivos que persiga una determinada evaluación y la delimitación de su propio objeto incidirán en las características formales del proceso. Como se ha visto en los capítulos anteriores existe una amplia variedad de objetivos y objetos de evaluación psicológica en los distintos modelos teóricos, y como consecuencia de ello se deriva cierta diversidad procesal. No obstante, se reconocen una serie de fases comunes a las mayoría de enfoques evaluativos, temática a la que se han dedicado abundantes y cuidadosos estudios (Fernández-Ballesteros, 1983c, 1992h, 1994d; 2004bc, Forns, 1990, 1993; Forns et al., 2002; Godoy y Silva, 1990, 1992; Silva, 1985, 1988, 1995c).

Como ya he señalado previamente la existencia de distintos objetivos de evaluación condicionarán, en cierto modo, procedimientos de actuación diferente. Por lo que en primer lugar, clarificaremos los distintos objetivos de la evaluación psicológica en la actualidad. Siguiendo a Fernández-Ballesteros (1992h) se pueden delimitar cuatro funciones diferentes de la Evaluación: Clasificación, Orientación, Selección y Tratamiento.

1) Clasificación: Ligada al ámbito clínico, la Evaluación Psicológica presentaba en sus inicios una meta exclusiva, la de clasificación. La asignación de un sujeto a una categoría del sistema clasificatorio psiquiátrico (psicótico, neurótico, retrasado mental...) ha sido y sigue

siendo para algunos psicólogos el único objetivo a alcanzar por un proceso de Psicodiagnóstico. Este diagnóstico puede ir acompañado de un pronóstico que aventurará cual será el comportamiento del sujeto en el futuro. Aunque la evaluación psicológica en la actualidad ha ampliado sus objetivos, en ocasiones un médico o un juez demandan la opinión del psicólogo sobre si el sujeto es clasificable en una u otra entidad nosológica, y en estas ocasiones se lleva a cabo este tipo de evaluación.

- 2) **Orientación o Consejo Psicológico:** Es aquella meta de la evaluación por la cual el estudio psicológico de un sujeto (un individuo, una pareja, un grupo, una institución) se realiza con el fin de dispensarle ayuda en orden a tomar decisiones o a establecer planes de acción referidos a su futuro. Estas decisiones pueden ser de orden personal o profesional. El proceso suele concluir con una orientación o consejo en relación al motivo de consulta planteado.
- 3) **Selección:** La evaluación puede tener por objetivo la selección cuando los informes recogidos sobre el sujeto o grupo de sujetos están dirigidos a elegir a la/s persona/s más apta/s en el desempeño de determinada actividad. La evaluación con fines de selección se opera tanto en el campo empresarial (selección de sujetos para determinados puestos de trabajo), como en el de la investigación (selección de sujetos para un estudio experimental). La selección de personal es un proceso cuyo objetivo final es establecer la idoneidad psicológica de una persona con referencia a un puesto de trabajo determinado. Este proceso exige el estudio de los puestos de trabajo, de los grupos en los que se va a integrar esa persona así como de las características de la organización. Por otra parte, dicho proceso requiere la determinación del perfil del puesto, la elección del modelo de selección adecuado y de los instrumentos mediante los que vamos a evaluar las diferencias individuales de los candidatos, una vez realizado el oportuno reclutamiento.
- 4) **Tratamiento o Cambio del comportamiento:** En este caso se requiere la evaluación de un sujeto (individuo, pareja, grupo, organización) con la finalidad de intervenir para producir cambios en sus conductas. En opinión de Fernández-Ballesteros (1992h) la diferencia entre los objetivos de orientación y los de tratamiento son, en muchos casos difíciles de establecer, ya que, en ocasiones, la orientación finaliza, precisamente, con la aplicación de un determinado tratamiento o intervención psicológica. Ejemplos de este tipo de objetivos son los siguientes: un escolar es evaluado con el fin de conocer qué tipo de intervención debe hacerse para incrementar su rendimiento escolar. Una pareja consulta con el fin de mejorar sus relaciones sexuales. En una institución de ancianos se realiza una evaluación para hallar cuales son las variables ambientales que están perturbando la conducta interpersonal de los residentes... Estos diferenciales objetivos de evaluación darán lugar a matices diferenciales en el procedimiento de actuación que se siga. Si los objetivos son el diagnóstico, la orientación, o la selección se empleará un enfoque correlacional mientras que si el objetivo es la intervención, el enfoque será el experimental. En buena medida la elección de un enfoque u otro dependerá de la propia demanda que se formule al evaluador y del objetivo evaluador. Evidentemente, no se procederá igual ante un caso de orientación profesional, que frente a una conducta fuertemente desestabilizada. De aquí, que en ocasiones, la dicotomía que suele establecerse entre evaluación tradicional versus

conductual venga impuesta por las propias metas evaluativas. De hecho, como veremos en el siguiente apartado, ambas perspectivas comparten fases comunes.

En la actualidad, el proceso de evaluación orientado a la toma de decisiones (Decision-Oriented Assessment, DOA) es concebido como una secuencia de pasos que guían al evaluador en la inevitable serie de decisiones que debe tomar antes de seleccionar la etiqueta diagnóstica más adecuada, dando respuesta a los requerimientos del cliente y a la orientación de caso (Westhoff, 2008).

5.3. FASES Y TAREAS DEL PROCESO DE EVALUACIÓN PSICOLÓGICA

Las distintas formas de relación entre evaluación y tratamiento están vinculadas a los presupuestos de análisis que sustentan las distintas escuelas psicológicas (Fernández-Ballesteros, 2004bc, Forns, 1993, Forns et al., 2002; Moreno, 2003; Muñoz, 2003; Pelechano, 1988c; Silva, 1995c). Así, la indicación de tratamiento puede ser vista de distinta forma según los modelos psicológicos básicos. Con la finalidad de clarificar las diferentes fases y tareas que se realizan durante el proceso desde diferentes marcos teóricos, a continuación se presentan las dos grandes perspectivas de evaluación:

1. La perspectiva correlacional: Proceso descriptivo-predictivo.
2. La perspectiva experimental: Proceso interventivo-valorativo.

5.3.1. LA PERSPECTIVA CORRELACIONAL: PROCESO DESCRIPTIVO-PREDICTIVO

La evaluación desde la perspectiva correlacional (tradicional) tiene por objetivo final transmitir la información diagnóstica al sujeto, a través de lo cual se orienta el caso. Habitualmente se considera que los procesos que responden a los modelos de evaluación tradicional (modelos psicométricos, psicoanalíticos, médico-psiquiátricos) se caracterizan por centrar su objetivo en la descripción, clasificación y/o predicción de la conducta. Tales procesos se basan primordialmente en metodologías correlacionales y pretenden evaluar rasgos o variables intrapsíquicas.

Muchos autores se han interesado por la sistematización del proceso diagnóstico desde una perspectiva clásica y/o tradicional. Investigadores teóricos y aplicados se ocupan de analizar de manera profunda esta cuestión, haciendo matizaciones que llegan a conformar una visión organizada de esta serie de tareas implicadas en el Psicodiagnóstico. Las opiniones de estos estudiosos del tema difieren en determinados aspectos, ya que mientras que unos aíslan un mayor número de fases, otros las agrupan estructurando un proceso más abreviado. Así mismo existen ciertas discrepancias respecto a la inclusión de algunas actividades concretas en determinada fase.

Con la finalidad de clarificar el proceso en la evaluación psicológica de enfoque correlacional, a continuación se presentan, las sistematizaciones planteadas por algunos autores contemporáneos.

SISTEMATIZACIÓN DE PELECHANO (1988)

Pelechano (1988c), desde su modelo de los parámetros, presenta un modelo de proceso evaluador en tres fases que intenta reflejar un procedimiento de actuación realista y eficaz. Este autor afirma que es difícil presentar un esquema general que sirva de guía en la realización de evaluaciones psicológicas debido a la heterogeneidad de los problemas tratados, los ámbitos de aplicación y las estrategias a utilizar, afirmando que hoy por hoy, no existe un esquema de trabajo común a todos los campos de la evaluación que no sea la defensa de una sensibilidad científica, aunque sugiere la existencia de líneas generales de actuación para determinadas parcelas de evaluación psicológica que sirvan como orientadoras a nivel formal (no en cuanto al contenido). Así, en el estudio del caso individual con vistas a un tratamiento distingue varias fases: análisis del problema, determinación de las posibles relaciones funcionales entre las situaciones vividas y respuestas del individuo, establecimiento de un posible modelo funcional, aplicación del tratamiento generado a partir de ese modelo de funcionamiento, seguimiento y control final de la eficacia.

En cuanto a los procedimientos de evaluación a utilizar, Pelechano recomienda en ese caso la utilización de estudios directos más que indirectos (empleo de observaciones, informes de terceras personas y autoinformes), mientras que en el estudio de los entornos recuerda el uso de cuestionarios y registros directos, afirmando que en el caso de la valoración de tratamientos y programas de actuación son varias las fuentes de información y criterios de bondad: desde los propios pacientes hasta los familiares e informes de los especialistas, cuyo contenido diferirá en función del tipo de aplicación. Así, partiendo de la idea de que no existe un único proceso dentro de la evaluación psicológica que permita aunar todos los tipos de procedimientos actuales propone tres fases (Pelechano, 1988c): 1) la primera delimitación del problema, 2) la recogida de información y 3) la integración-sistematización de la información recogida.

- 1. Delimitación del problema:** Esta fase implica una serie de decisiones por parte del evaluador: (a) Decidir si se trata de un problema que cae dentro de su competencia o no; (b) Determinar el peso relativo que el evaluador debe dar a las primeras informaciones que obtiene en un caso concreto; y (c) Especificar los problemas en términos operativos, determinando los tipos principales de subproblemas y promoviendo la elección de la adecuada estrategia de evaluación, el método y los procedimientos más adecuados, así como cuales son las informaciones que deben recogerse.
- 2. Recogida de Información:** Con esta finalidad se utilizan procedimientos y pruebas que deben de cumplir cuatro requisitos: (a) Respetar a las persona/s evaluada/s debiendo de pedir su consentimiento para la realización de grabaciones de vídeo o audio; (b) Los procedimientos de recogida de información no deberán molestar a la persona evaluada ni someterla a presión física ni psicológica; (c) Los procedimientos de recogida de información no deberán interferir o interferirán al mínimo con el proceso comportamental; (d) La selección de la información y la instrumentación se realizará en función de su bondad relativa a los objetivos perseguidos. Si se utilizan instrumentos psicométricos deberá tenerse en cuenta que su fiabilidad y validez sea adecuada al estudio que se realiza. En cuanto a los contenidos a recoger plantea el uso de estrategias generales (selección de entrevista, tests y

técnicas psicológicas adecuadas así como la observación directa preferentemente mecanizada) y estrategias específicas (sub-evaluaciones que pueden surgir dentro del proceso general de evaluación cuando al estudiar el problema se detectan otros que requieren ser objeto de estudio).

3. Integración-Sistematización de la información: En esta fase se responde a las preguntas que se plantearon en la primera delimitación del problema y que fueron el motivo de la evaluación. Después de recoger la información debe integrarse y plasmarse en un informe. Este (a) debe de describir un modelo de funcionamiento bien sea individual, colectivo u institucional; (b) tendrá en cuenta a quién va dirigido; (c) no se derivarán de él consecuencias negativas para la persona evaluada; y (d) debe de ser fundamentalmente descriptivo, ofreciendo los datos de modo ordenado y referidos a los objetivos previamente explicitados, debiendo restringirse al máximo la interpretación inferencial. Además, deben incluirse recomendaciones y o planes de intervención lo más claros y eficaces posibles.

Para Pelechano (1988c) una descripción psicológica debe de cumplir tres requisitos: Normalización, Validación y Calibración. Por normalización se refiere a los métodos y procedimientos seguidos en la obtención de información, su análisis y guía interpretativa deben ser claros como para que cualquier evaluador alcance los mismos resultados, lo que implica la objetivación de las situaciones, operacionalización de las variables evaluadas y objetividad en las técnicas y procedimientos de análisis de la información recogida. Por validación se refiere a la validez de constructo, es decir, que exista una correspondencia entre la categoría de evaluación que se está utilizando y los correlatos de respuesta. Y calibración que hace referencia a la existencia de medidas, escalas, puntuaciones e indicadores cuantitativos. Así, después de delimitar la existencia de una categoría diagnóstica es importante conocer el grado en que está presente.

En síntesis, para este autor (Pelechano, 1988c) el proceso de evaluación es un proceso interactivo entre evaluador y evaluado(s). En este proceso pueden distinguirse tres momentos: delimitación del problema, recogida de información y oferta de respuestas para los problemas planteados. Dentro de este proceso de evaluación se incorporan aspectos psicométricos tradicionales. Estos aspectos, sin embargo, no agotan las posibilidades dado que pueden incorporarse, aspectos cualitativos no susceptibles de medida directa, estimaciones y estadística intensiva. En su opinión el proceso de evaluación psicológica no se encuentra adscrito a ninguna teoría o escuela concreta de pensamiento psicológico. El profesional encargado de realizar la evaluación ha de conocer y saber utilizar los elementos teóricos que se encuentran apoyando las técnicas, tests, pruebas o procedimientos de recogida e interpretación de la información recogida, elaborando, a la par que va recogiendo la información, el modelo funcional que parece irse ajustando a los resultados obtenidos.

SISTEMATIZACIÓN DE RODRÍGUEZ GONZÁLEZ (1989)

Rodríguez González (1989) analizando distintas estructuraciones, y en un intento de ordenar las fases del proceso diagnóstico propone 6 fases:

- 1. Información Preliminar:** Fase en la que se recoge información necesaria, se operativiza el motivo de consulta, y se decide si la actuación del psicólogo es necesaria. Se emplea para ello la observación, la entrevista con el cliente y con personas que pueden verse implicadas, así como la historia clínica.
- 2. Elaboración de Hipótesis:** Fase en la que se elabora una teoría del caso. El marco teórico referencial (psicometrista, médico-psiquiátrico...) del psicólogo condicionará el tipo de hipótesis que plantee.
- 3. Diseño de la Evaluación:** En esta fase se seleccionan las técnicas o tests que permitan recoger información. En este aspecto existen discrepancias entre autores de distinto enfoque que plantean el uso de técnicas globales o focalizadas. En el contexto de esta discrepancia y en una posición integradora, este autor propone un empleo inicial de técnicas globales que faciliten una visión genérica del individuo y posteriormente un empleo de técnicas focalizadas con las que explorar en profundidad los problemas concretos. Se emplean la Entrevista, la Observación y los Tests Psicológicos.
- 4. Exploración:** En esta fase se aplican y corrigen los instrumentos y técnicas necesarias para recoger información relevante al caso.
- 5. Interpretación y procesamiento de los datos:** Los datos obtenidos mediante la aplicación de diversos instrumentos nos proporcionan una imagen de la persona evaluada, que debe ser contrastada con las hipótesis planteadas. Ahora se valora si la información disponible es suficiente para dar respuestas a las suposiciones iniciales. Si no fuera así se aplican nuevos instrumentos, pero si se dispone de la información necesaria se emite un Juicio Diagnóstico. También en este aspecto se constatan discrepancias, ya que dependiendo de la formación del psicólogo debe elegir un enfoque clínico o estadístico. Mediante el enfoque clínico se hacen las inferencias necesarias que permitan pasar de hechos conocidos a otros que se suponen a partir de los iniciales. En el caso de la orientación estadística o actuarial, el informe se basa en procedimientos matemáticos y bases de datos. Desde una perspectiva integradora Rodríguez González (1989) plantea la combinación de ambos.
- 6. Comunicación de los resultados:** En esta fase se elabora un informe en el que se registran los resultados, las conclusiones y recomendaciones planteadas al caso, que podrá ser comunicado en forma verbal o escrita. En este enfoque tradicional el evaluador no realiza la terapia por lo que su informe será de gran importancia para el profesional que la realice. El informe incluye los datos de identificación del paciente, el motivo de la consulta, los instrumentos y técnicas empleadas así como sus resultados, conclusiones y recomendaciones.

Al hilo de esta fase de devolución de los resultados al propio paciente, a sus allegados u otros, resulta necesario discriminar el tipo de información que es necesario comentar con cada uno de ellos. Este aspecto ha promovido multitud de dudas: ¿Se le debe dar toda la información al enfermo?, ¿Hasta qué punto se deben proporcionar datos a la familia?, ¿Qué otras personas tienen acceso a la información vía oral o vía documento escrito?... Todos estos interrogantes son hechos en los que un proceder ético juega un papel esencial. Si bien el Código

Deontológico del Colegio Oficial de Psicólogos subraya el secreto profesional y el derecho a la intimidad, su carácter general no facilita soluciones directas a problemas individuales. Ante esta situación, el psicólogo tiene que guiarse de sus criterios y responsabilidades.

Otro problema que de forma paulatina se nos va presentando de manera especial a los evaluadores lo constituyen las peticiones, cada vez más frecuentes, de actuación en Peritajes Judiciales. La consideración del problema deviene a partir de las repercusiones que estos requerimientos tuvieron, ya hace años, en otros países para el colectivo de Psicólogos.

LA SISTEMATIZACIÓN DE FORNS (1993)

El proceso de evaluación psicológica infantil según Forns (1993) debe basarse en la concepción dinámica, interaccional y global de la conducta. Propone un proceso de evaluación como un “sucesivo encadenamiento de decisiones”; es decir, establecer la relación entre conductas y situaciones y los elementos que han mantenido la conducta problema. Para ello sugiere un análisis diacrónico del objetivo de estudio. Las tareas de cada fase del proceso tienen distinta amplitud en función de sus objetivos. La práctica de la evaluación tradicional, correlacional en la sistematización de Forns (1993) se estructura en tres fases:

1. Fase de determinación de las bases para la exploración:

- Planteamiento del problema.
- Búsqueda selectiva y análisis de la información relevante para captar sus variables.
- Establecimiento de hipótesis diagnósticas.
- Elaboración de un proyecto de exploración que conlleva la elección de técnicas de exploración adecuadas a través de las cuales contrastar las hipótesis, valorar el problema y planificar las sesiones de exploración.

2. Fase de exploración psicológica:

- Valoración del estado motivacional del individuo.
- Valoración de las condiciones situacionales no-controlables que puedan condicionar los datos obtenidos en el examen.
- Aplicación de las pruebas y análisis de los datos obtenidos articulando los datos cualitativos y cuantitativos.

3. Fase diagnóstica y de toma de decisiones:

- El análisis de los datos obtenidos permite contrastar las hipótesis planteadas, y formular un diagnóstico nosológico.
- Toma de decisiones a cerca del futuro del problema que supone plantear un pronóstico y formular indicaciones del tratamiento más adecuado, que puede ser de naturaleza muy diversa (orientación escolar o académica, asesoramiento, exploración médica, derivación a tratamiento psicológico...).

LA SISTEMATIZACIÓN DE FERNÁNDEZ-BALLESTEROS (2004bc)

Para Fernández-Ballesteros (2004b) una vía para llegar al estudio científico de un sujeto desde la Psicología consiste en abordar éste recogiendo la información pertinente a través de una serie de técnicas, tests o instrumentos, por medio de los cuales cumplir los objetivos del caso, llegando a describir, clasificar, seleccionar, orientar o recomendar un tratamiento. Este proceso de evaluación en su opinión se lleva a cabo mediante 4 fases:

- 1. Primera recogida de Información sobre el caso:** Durante esta fase resulta necesario recabar datos sobre los siguientes aspectos: 1) Especificar la demanda y fijar objetivos sobre el caso (indagación de la razón por la que se solicita evaluación, sobre lo que se desea conseguir con ella); y 2) Especificar las condiciones históricas y actuales potencialmente relevantes (historia del sujeto, condiciones actuales socio-ambientales y biológicas).
- 2. Formulación de hipótesis y deducción de enunciados verificables:** En función de las observaciones e informaciones recibidas se formulan hipótesis sobre el caso. Las hipótesis deben ser contrastables y estar basadas en los datos empíricos disponibles. En este momento influirán los conocimientos que tenga el evaluador sobre las funciones psicológicas, las covariaciones bien establecidas entre clases de respuesta, los fundamentos fisiológicos y neurológicos de la conducta, las variables ambientales que mantienen o controlan las conductas problema... En esta fase son 4 los tipos de supuestos que pueden ser formulados: *De Cuantificación* (constatación de la presencia y medida de la conducta objeto de estudio), *De semejanza* (constatación de relaciones de semejanza con otros sujetos clasificados en determinada categoría), *De Asociación Predictiva* (realización de predicciones apoyadas en asociaciones contrastadas empíricamente), y *De Relación Funcional o Explicativa* (formulación de causas a las que se debe la afección del sujeto).
- 3. Contrastación a través de la administración de tests y otras técnicas de evaluación:** Las actividades que se llevan a cabo en la tercera fase del proceso evaluador son la aplicación de las técnicas a través del procedimiento establecido y la elaboración de los resultados en orden a la comprobación de las hipótesis. Los resultados obtenidos mediante la aplicación de técnicas permitirán comprobar las hipótesis planteadas respecto al caso.
- 4. Comunicación de los resultados: el informe.** En esta fase se establecen las conclusiones, se describe y clasifica a los sujetos, operándose predicciones y tomando decisiones respecto al caso planteado, como por ejemplo, la propuesta de un tratamiento... En esta fase se integran los resultados, y posteriormente, se elabora un informe y se comunican los resultados del proceso al sujeto o a sus allegados.

Como señalaba al principio, al presentar un esquema de evaluación desde la perspectiva correlacional (descriptivo-predictivo) se incurre necesariamente en reduccionismos que en ocasiones no permiten clarificar los matices diferenciales, entre psicometristas o psicoanalistas, que se dan en las distintas perspectivas ubicadas dentro de este epígrafe. Aunque ambas perspectivas etiquetadas como tradicionales pueden compartir los presupuestos previamente señalados.

5.3.2. LA PERSPECTIVA EXPERIMENTAL: PROCESO INTERVENTIVO-VALORATIVO

En la perspectiva experimental se considera que existe un punto de articulación entre evaluación y tratamiento. Entre evaluación y tratamiento existe una estrecha relación y feedback continuo. La evaluación desde esta perspectiva se plantea cuando la finalidad es operar un cambio en el comportamiento, una intervención psicológica. Ahora el evaluador debe planificar el tratamiento en función de una evaluación previa, y posteriormente valorar los efectos del mismo. El proceso experimental de la Evaluación y Diagnóstico Psicológicos, es en su primera parte idéntico al proceso correlacional (descriptivo) y será a partir de los resultados de la evaluación previa cuando el evaluador deba plantearse hipótesis funcionales que verificará a través de la manipulación experimental, mediante un tratamiento de una serie de variables relevantes al problema (independientes) y una comprobación de sus efectos sobre esas conductas-problema (variables dependientes).

LA SISTEMATIZACIÓN DE ECHEBURUA (1993)

Echeburua (1993a) después de analizar diversos estudios que sistematizan las fases del proceso en evaluación-intervención conductual propone un proceso sintético que se organiza en 5 fases:

- 1. Identificación de las conductas problemáticas y de las variables controladoras** (*análisis topográfico y funcional de las conductas-problema*). *Evaluación de los recursos conductuales y la motivación para el tratamiento*. Para esta evaluación se emplean métodos directos (observación, auto-observación, registros psicofisiológicos) e indirectos (autoinforme, entrevista).
- 2. Formulación de hipótesis sobre los trastornos de la conducta en base a los resultados obtenidos en la fase previa:** Se plantean hipótesis sobre el origen, el mantenimiento y la modificación de cada una de las conductas-problema.
- 3. Establecimiento de los objetivos de intervención y establecimiento del orden de intervención.**
- 4. Selección de los procedimientos terapéuticos y puesta en práctica del programa de intervención.**
- 5. Evaluación de la terapia:** Evaluación intratratamiento, postratamiento, y de seguimiento.

SISTEMATIZACIÓN DE FERNÁNDEZ-BALLESTEROS (1992h, 1994d, 2004c)

Para Fernández-Ballesteros (1992h) las 4 primeras fases del proceso experimental son coincidentes con las del proceso correlacional. Es a partir de la quinta fase, una vez formuladas las hipótesis funcionales, cuando comienza, propiamente, el momento experimental (interventivo-valorativo) del proceso. Las fases que la autora plantea en esta obra son básicamente las mismas que las planteadas posteriormente (Fernández-Ballesteros, 1994d, 2004c) en las que aborda las fases del proceso en evaluación conductual, estableciendo de

este modo una analogía entre ambos conceptos. En su opinión, el proceso experimental del proceso diagnóstico se configura en 9 fases:

- 1. Primera recogida de información: especificación de la demanda del problema:** En esta fase se delimita el problema por el que se viene a consulta, recogiendo información sobre las condiciones ambientales, personales y/o biológicas, pasadas y actuales, relevantes.
- 2. Primera formulación de hipótesis y deducción de enunciados verificables:** Después de verificar la existencia de un problema, se analizan las condiciones ambientales, personales o biológicas asociadas al problema, se formulan las hipótesis del caso. El establecimiento de las mismas dependerá de la evidencia empírica que, procedente de las teorías sobre el problema tenga el evaluador. En base a las hipótesis el evaluador seleccionará los instrumentos de medida adecuados para explorar las variables objeto de estudio.
- 3. Contratación inicial:** Se trata de verificar las predicciones establecidas. Se recoge información aplicando las técnicas y tests seleccionados con la finalidad de valorar las variables objeto de análisis referidas al ambiente, a la persona y al organismo. En esta fase, el dominio del evaluador relativo a los instrumentos tendrá gran relevancia.
- 4. Resultados Iniciales:** Una vez analizados los datos, se suministra al sujeto la información relevante, pero ello se opera en forma oral, en el transcurso de una entrevista.
- 5. Formulación de hipótesis funcionales:** En base a los resultados obtenidos se plantean una serie de supuestos funcionales explicativos sobre el caso. Como señalan Haynes, Uchigakiuchi, Meyer, Orimoto y Blaine (1993) a la hora de establecer modelos causales, se asigna un peso hipotético a cada una de las variables explicativas. La asignación de ponderaciones a las variables explicativas es una tarea fundamentalmente subjetiva del evaluador que matemáticamente puede ser resuelta mediante la teoría de la decisión, más que aplicando coeficientes de regresión (Fernández-Ballesteros, 1994d). Con la finalidad de plantear los supuestos funcionales explicativos del caso se realizan tres tareas: (1) establecimiento de los dispositivos de medida que actuarán como variables dependientes y de los criterios de cambio; (2) selección y medición de la/s variable/s independiente/s a manipular; y (3) selección de la/s técnica/s de manipulación de la/s variable/s independiente/s o del paquete de tratamiento.
- 6. Tratamiento: Recogida de datos pertinentes a las hipótesis y aplicación del tratamiento:** Por tratamiento se entiende cualquier forma de intervención psicológica que tenga por objetivo un cambio de la/s actividad/es psicológica/s del sujeto. Antes de aplicar el tratamiento el psicólogo deberá establecer: (1) el diseño de intervención y valoración a través del cual serán contrastadas las hipótesis; (2) la selección y aplicación de las técnicas de evaluación y control de posibles variables intervinientes y contaminadoras; y (3) en ocasiones, la evaluación continua del progreso del tratamiento.
- 7. Contratación de hipótesis y valoración del tratamiento:** En esta fase el evaluador realiza: (1) una nueva evaluación de las conductas-problema con el fin de comprobar si se

han alcanzado las metas terapéuticas propuestas; (2) una constatación de si los datos avalan las hipótesis inicialmente formuladas concluyendo con la información al sujeto sobre todo ello; y (3) la constatación de si el tratamiento ha sido efectivo.

8. **Resultados:** El proceso de evaluación termina informando al sujeto y/o sus allegados de los resultados obtenidos, en ocasiones de forma verbal y en otras en forma de informe escrito.
9. **Seguimiento:** El seguimiento implica una nueva evaluación transcurrido un tiempo determinado desde la aplicación del tratamiento, a fin de comprobar si se mantienen los resultados obtenidos.

Respecto a las técnicas que se emplean según la etapa del proceso de Evaluación en la que se esté operando, Fernández-Ballesteros (1992h; 1994d) plantea distintos tipos de técnicas clasificadas en tres niveles relacionados con el coste (bajo-alto) y con la fase del proceso (1-9):

- Primer Nivel: Técnicas de amplio espectro (entrevista, autobiografía, listados de conductas, escalas de apreciación, observación asistemática, datos de archivo, rastros...).
- Segundo Nivel: Técnicas generales (autorregistros, observación sistemática, pruebas estandarizadas de ejecución...) y Técnicas específicas (cuestionarios o escalas de autoevaluación, escalas de apreciación de otros).
- Tercer Nivel: Códigos de observación sistemática (expertos, allegados...) y Técnicas objetivas (registros fisiológicos...).

SISTEMATIZACIÓN DE FORNS, ABAD, AMADOR, KIRCHNER Y ROIG (2002)

Forns et al. (2002) subrayan las siguientes fases en el proceso de evaluación psicológica:

1. Responder a una petición de exploración psicológica y obtener información significativa.
2. Reformular la petición en términos psicológicos y formular hipótesis sobre la naturaleza de la consulta.
3. Recoger información específica para verificar las hipótesis de presunción diagnóstica: el examen o exploración psicológica.
4. Establecer el diagnóstico psicológico.
5. Transmitir la información debidamente conceptualizada seleccionada y organizada al destinatario pertinente.
6. Seleccionar y planificar las estrategias de tratamiento.
7. Aplicación del tratamiento.
8. Valoración del cambio producido como fruto del tratamiento.
9. Seguimiento o control sucesivo del establecimiento del cambio.

SISTEMATIZACIÓN DE MUÑOZ (2003)

El proceso de evaluación psicológica desde el ámbito clínico y de la salud de Muñoz (2003) denominado proceso de Evaluación Psicológica Clínica (EPC), comprende dos momentos diferenciados: evaluación y tratamiento. La evaluación está compuesta por tres ejes simultáneos de actividad: 1) Análisis descriptivo, 2) Análisis funcional y formulación clínica, y 3) Diagnóstico. El análisis descriptivo incluye la exploración inicial, identificación de problemas, secuencias de cada problema, desarrollo de cada problema y esquema narrativo del caso, variables de la persona y del contexto. El análisis funcional y formulación clínica incluye las primeras hipótesis, las hipótesis de trabajo. En el eje del diagnóstico encontramos la impresión diagnóstica, el diagnóstico (DSM-IV-TR; ICD-10), criterios para cada trastorno, datos epidemiológicos y modelos teóricos y psicopatológicos de trastornos. El tratamiento se divide en: 1) Diseño del tratamiento, 2) Evaluación continuada del cambio durante el tratamiento, 3) Evaluación de la efectividad de la intervención.

SISTEMATIZACIÓN DE MORENO (2003)

Moreno (2003), tomando como referencia los trabajos de Fernández-Ballesteros (1992), enuncia 9 fases.

- 1. Primera recogida de información: especificación de la demanda y del problema**
 - 1.1. Especificar la demanda
 - 1.2. Fijar objetivos sobre el caso
 - 1.3. Establecer las condiciones históricas y actuales potencialmente relevantes
- 2. Primera formulación de hipótesis y deducción de enunciados verificables**
 - 2.1. Plantear supuestos (momento inductivo)
 - 2.2. Plantear deducciones verificables sobre el caso (momento deductivo)
- 3. Contrastación inicial de hipótesis**
 - 3.1. Contrastar hipótesis y enunciados formulados en la fase 2
- 4. Resultados**
 - 4.1. Comunicar los resultados del proceso evaluativo
- 5. Formulación de hipótesis funcionales**
 - 5.1. Establecer los supuestos funcionales sobre el caso
 - 5.2. Deducir predicciones verificables
- 6. Recogida de datos pertinentes a las hipótesis funcionales**
 - 6.1. Planificación del tratamiento
 - 6.2. Aplicación del tratamiento
- 7. Valoración de resultados**
 - 7.1. Segunda evaluación de las conductas-problema y de las metas establecidas
 - 7.2. Comprobar si se han alcanzado las medidas de las variables relevantes
 - 7.3. Comprobar si los datos avalan las hipótesis funcionales formuladas y valoración del tratamiento
- 8. Resultados**
 - 8.1. Comunicar los resultados obtenidos después de la aplicación del tratamiento

9. Seguimiento

- 9.1. Nueva recogida de datos
- 9.2. Comprobar si se mantienen los resultados del tratamiento

Fase 1: Primera recogida de información: especificación de la demanda y del problema

Recopilar información con un objetivo doble: especificar en términos operativos el motivo de la demanda y los objetivos de la consulta, y obtener los datos sobre las condiciones pasadas y actuales potencialmente relevantes. La primera toma de contacto se inicia con la entrevista psicológica con el objetivo de averiguar el motivo de la consulta, las conductas problema principales y los objetivos o expectativas del cliente. Se trata del análisis de las demandas, quejas y metas. Si se trata de una evaluación psicológica infantil, la demanda es formulada casi siempre a través de terceros; principalmente, los padres. Los objetivos que se pueden solicitar son los de descripción, clasificación, predicción y explicación del comportamiento con la finalidad de orientar o planificar un tratamiento y modificar la conducta problema. Conocido el motivo de la consulta, es necesario recabar información acerca de las condiciones actuales y pasadas de la persona (autobiografía en adultos, padres que informan en niños). La información obtenida con la entrevista puede completarse con otras técnicas y otras fuentes, además del niño. En la evaluación psicológica infantil también debe haber entrevista con el pequeño, intentando crear un ambiente agradable, escoger el mejor momento del día, etc. Por otro lado, las variables ambientales son muy importantes ya que el niño es muy permeable a la influencia social, y, por otro lado, los niños son diferentes según los ambientes. A veces, puede ser interesante conocer las interacciones padres-hijo para conocer los estilos interactivos, refuerzos... No olvidar la influencia de algunos problemas de los adultos sobre el comportamiento de los más jóvenes. Es necesario un buen conocimiento del desarrollo infantil, tener habilidades sociales y experiencia en el trato con niños y padres y conocer los instrumentos de evaluación utilizados.

Fase 2: Primera formulación de hipótesis y deducción de enunciados verificables

Incluye dos momentos: inductivo (formulación de hipótesis) y deductivo (establecimiento de enunciados o predicciones verificables o cuantificables). La información recopilada en la fase anterior sirve para conseguir los objetivos de esta fase. Existen cuatro tipos de supuestos o hipótesis sobre un caso: cuantificación, semejanza, asociación predictiva y relación funcional o explicativa. Los supuestos de cuantificación se formularán en casos que requieran verificar la frecuencia con la que se da una determinada conducta. Los de semejanza cuando queramos comparar una determinada conducta frente a otras de otras personas. La asociación predictiva presupone relaciones entre antecedentes y consecuentes de la conducta. Se utilizarán supuestos de relación funcional o explicativa cuando la conducta puede relacionarse con aspectos funcionales. Con el fin de comprobar las hipótesis o suposiciones, se establecen enunciados que pueden ser verificables. Para ello se parte del análisis de las variables relevantes asociadas al problema, y de la selección de las técnicas o instrumentos que servirán de base para deducir los enunciados.

En cuanto a la evaluación psicológica infantil, una vez formuladas las hipótesis del caso, las variables que se deberán escoger deben estar claramente especificadas y referirse a

distintos aspectos del funcionamiento general del niño. Los instrumentos han de ser más específicos que los utilizados en la fase anterior. Han de evaluar específicamente el problema como otros aspectos que puedan estar relacionados directa o indirectamente con el mismo. Utilizar pruebas autoaplicadas y heteroaplicadas. Esto permitirá realizar predicciones verificables sobre el caso. También es imprescindible utilizar distintos informadores. En la evaluación infantil hay unanimidad en la utilización de la evaluación “multicomponente” (Silva, 1995), que analice distintas variables (personales, ambientales, biológicas) a través de distintos instrumentos y utilizando para ello distintas fuentes de información. Esta perspectiva ecológica y plural de la evaluación infantil está apoyada por la mayoría de especialistas del área.

Fase 3: Contrastación inicial de hipótesis

Verificar las predicciones establecidas para lo cual se aplican las técnicas de evaluación elegidas anteriormente para recoger información pertinente a los supuestos. Las tres tareas que se realizan son: Preparación del material y del procedimiento de administración. Los autorregistros y registros observacionales específicos requieren más tiempo de preparación. Las pruebas estandarizadas es preciso conocerlas (administración, corrección, interpretación). En cuanto al procedimiento de administración hay que tener en cuenta el orden de aplicación y el tiempo necesario. Suelen ser necesarias varias sesiones y en los niños es preferible por la mañana. Preparación del demandante y, en su caso, de sus allegados implicados en la evaluación. Con el objetivo de conseguir una actitud o conducta activa frente a la evaluación, fundamental en el proceso infantil (enseñar a observar, manuales para padres o maestros. Es preciso motivarles para que participen. La aplicación de los procedimientos de recogida de información establecidos para recoger la información pertinente. Es el grueso de la evaluación, se deben usar varias sesiones para recoger la información necesaria para contrastar las hipótesis y enunciados formulados anteriormente.

En cuanto a la evaluación infantil se ha de empezar por la evaluación a otras fuentes y el tiempo que utilizan para registrar la conducta del niño en distintos contextos se deberá aprovechar para evaluar al niño. Es necesario atender a su edad, signos de cansancio... Se recomienda la evaluación de los aspectos instrumentales e intelectuales del niño. La exploración ha de ser individualizada preferentemente, aunque a veces debido a la corta edad puede ser de ayuda la presencia de la madre. La tarea de corrección, valoración y análisis de la información obtenida puede incorporarse al final de esta fase o al principio de la siguiente. El incluirla aquí es debido a que el objetivo de esta fase es comprobar los supuestos o hipótesis formulados en la segunda fase. Si no se verifican los enunciados, habrá que volver a la fase anterior y reformular las hipótesis.

Fase 4: Resultados

Se informa de los resultados obtenidos a partir de la exploración realizada en la fase anterior. Se lleva a cabo en una entrevista en la que se dan respuesta a los objetivos planteados en la primera fase. Así, deberá llevar el diagnóstico del caso y la planificación de actuaciones. El informe psicológico escrito sobre la evaluación infantil debe expresar el problema y no emitir, según los casos, conclusiones definitivas e inamovibles. Con el informe oral se realizarán las pertinentes aclaraciones, subsanar dificultades, etc. Si el objetivo de la

demanda es el de diagnóstico, orientación o selección, el proceso psicológico terminará en este momento. Pero, si nos solicitan además un programa de intervención psicológica, entonces del proceso de evaluación psicológica continuará hasta la novena fase.

Fase 5: Formulación de hipótesis funcionales

A partir de los resultados emitidos en la fase anterior se formulan nuevos supuestos funcionales o hipótesis explicativas sobre el caso, lo que exige establecer enunciados o predicciones verificables, que serán contrastados a través de un diseño experimental en el que se introducirá el tratamiento (fase 6) y su valoración (fase 7). Los principales objetivos de esta fase son dos: 1) establecer los supuestos funcionales sobre el caso: selección de las conductas-problema a modificar y de las variables relevantes a manipular y de las técnicas de manipulación; 2) deducir predicciones verificables. Para conseguir estos objetivos se deben realizar las siguientes tareas:

1. *Selección de las variables dependientes y de los criterios de cambio u objetivos.* Seleccionar las conductas que deben ser modificadas y ordenarlas para su tratamiento. Una vez determinadas las conductas objetivo deberán operativizarse a través de unas medidas concretas, con el fin de observar de forma objetiva los cambios que se produzcan.
2. *Selección de las variables independientes.* Se seleccionan las variables explicativas del caso, las que mantienen, controlan o explican el problema. Tanto las variables personales, ambientales como las biológicas pueden explicar la conducta problema (multicausalidad de la conducta). Las únicas manipulables son las variables personales y ambientales, las biológicas necesitan la intervención de otros profesionales.
3. *Selección de las técnicas de manipulación de las variables independientes.* Elegir técnicas efectivas, que respondan a las características de personalidad. En el área infantil, la modificación de conducta ha demostrado ofrecer alternativas de solución.
4. *Elaboración de un modelo explicativo.* Formalizar un modelo funcional; es decir, establecer asociaciones entre las variables independientes y las dependientes.
5. *Deducción de predicciones.* La consecución de las tareas anteriores permite la deducción de predicciones sobre los cambios que supuestamente se darán en la persona una vez aplicado el programa de intervención. Estas predicciones se verificarán en la fase siguiente.

Fase 6: Recogida de datos pertinentes a las hipótesis funcionales

Esta fase el objetivo central es la aplicación del tratamiento. Los objetivos principales son: 1) establecimiento del diseño de intervención y valoración; 2) control de posibles variables potencialmente contaminadoras; y 3) aplicación del tratamiento. Para llevar a cabo estos objetivos se proponen las siguientes seis tareas:

1. *Elección del diseño adecuado.* El profesional debe conocer los posibles diseños experimentales de series temporales, caso único o de replicación intrasujeto. El registro que se lleva a cabo en el pre-tratamiento, es decir, en ausencia de tratamiento se denomina “línea base”, mientras que la evaluación que se realiza en el post-tratamiento, es decir después de aplicar la intervención es la “fase experimental”.
2. *Selección y aplicación de las técnicas de evaluación y control de posibles variables contaminadoras de los efectos del tratamiento.* Es necesario controlar que los datos procedentes del tratamiento estén libres de errores y sesgos. Las fuentes de error pueden ser variadas. Las expectativas de la persona frente al tratamiento pueden ejercer un papel modulador, para lo cual se propone la evaluación y control del nivel motivacional antes de iniciar la intervención. En el caso infantil, como son los padres los que tienen expectativas acerca del tratamiento, aunque éste se le aplica al niño, hay que evaluar el nivel motivacional tanto del niño como de los padres.
3. *Preparación del demandante y, en su caso, del cliente.* Ahora se trata de prepararles para la aplicación del tratamiento. Los padres participan en muchas de las intervenciones aplicadas a los hijos, por lo que es muy importante su preparación para evitar cualquier fuente de error.
4. *Aplicación del tratamiento.* Es el grueso de esta fase, es la aplicación de la técnica o técnicas de manipulación de las variables independientes. En su caso, evaluación continua del progreso del tratamiento. A veces, durante el tratamiento se sigue evaluando todas o algunas de las conductas que se pretenden modificar con el fin de tener una información continua de los efectos del tratamiento. En el caso de no producirse los cambios se puede introducir alguna modificación o interrumpir el tratamiento.
5. *Finalización del tratamiento.* Puede darse de distintas formas: conseguir los objetivos terapéuticos, terminar el plazo establecido para los cambios en las variables dependientes, y por voluntad del demandante, en el caso de evaluación psicológica infantil, suelen ser los padres del niño.

Fase 7: Valoración de resultados

Los objetivos de esta fase son: 1) evaluación de las conductas-problema y de las metas establecidas; 2) comprobación de si se han alcanzado las medidas de las variables relevantes y si los datos avalan las hipótesis funcionales formuladas; y 3) valoración del tratamiento. Para conseguir estos objetivos es preciso realizar las siguientes tareas:

Aplicación de las técnicas de medida de las conductas problema y/o de las metas de tratamiento y, en su caso, de las variables relevantes. Evaluación de las variables dependientes o conductas objetivo. Los instrumentos serán los mismos y los resultados servirán para comprobar el modelo causal.

Análisis de datos. Para comprobar los cambios producidos en las variables dependientes a partir de la manipulación de las independientes. Hay dos formas de análisis de datos: a partir de un análisis estadístico y a partir de la interpretación gráfica de los datos.

Comprobación de las hipótesis formuladas y valoración del tratamiento. Los resultados obtenidos nos permiten comprobar las hipótesis funcionales formuladas y la efectividad del tratamiento aplicado. La verificación ha de realizarse en dos direcciones: 1) Comprobar la relación funcional entre la variable dependiente e independiente; y 2) Comprobar si existe relación funcional entre la variable independiente y un determinado tratamiento. Si no se hubieran confirmado los supuestos formulados o se hubieran producido otros trastornos, sería necesario volver a la fase 5 (Formulación de hipótesis funcionales) o en su defecto a la fase 2 (primera formulación de hipótesis y deducción de enunciados verificables).

Fase 8: Resultados

El objetivo es comunicar los resultados obtenidos después de la aplicación del tratamiento, a través de un informe oral y/o escrito de los resultados obtenidos. El informe en evaluación psicológica infantil debe contemplar además de los cambios post-tratamiento, otros posibles cambios evolutivos que puedan afectar a los resultados de la intervención.

Fase 9: Seguimiento

No sólo es importante que el tratamiento tenga efecto sino que también lo es que no se desvanezca con el paso del tiempo. Los objetivos de esta fase son la nueva recogida de datos y la comprobación de si se mantienen los resultados del tratamiento. Para la realización de estos objetivos, es necesario llevar a cabo algunas de las tareas ya realizadas en distintos momentos del proceso: a) Nueva toma de contacto con el cliente (fase 1); b) Nueva evaluación de las conductas objetivo (fase 3 y 7); c) Nuevo análisis de datos, comparándolos con los obtenidos antes e inmediatamente después del tratamiento (fase 7); d) Comprobación de la estabilidad y mantenimiento de los resultados del tratamiento y, en su caso, informe al cliente y sus allegados (fase 7).

5.3.3. REFLEXIONES SOBRE LA EVALUACIÓN PSICOLÓGICA APLICADA COMO PROCESO CORRELACIONAL VERSUS EXPERIMENTAL

Desde los enfoques conductistas se ha considerado a la evaluación tradicional de escaso rigor científico. Ciertamente la metodología experimental dota a todo el proceso de evaluación de mayor rigor en la formulación de objetivos de análisis. Elige un procedimiento de orden hipotético-deductivo evitando inferencias de alto grado, pretende mayor precisión en la medida y establece mayor control, siendo sus resultados verificables. Cabe, no obstante, señalar que dentro de la evaluación tradicional se incluyen enfoques que difieren en grado de precisión y verificabilidad. Sin embargo, podríamos preguntarnos si en la práctica clínica cotidiana es posible llevar a cabo las pautas de la metodología científica.

En este sentido juzgo interesante la opinión de Pelechano (1988c) quien considera que la evaluación psicológica debe seguir los dictados de la metodología científica, de manera

rigurosa cuando trata de producir un conocimiento básico, pero no debe confundirse con ella en su vertiente aplicada. Por lo que se refiere a la consideración aplicada de la evaluación psicológica, los pasos, fases y requisitos de la investigación científica representan un modelo regulador al que hay que acercarse dentro de lo posible pero, de ninguna manera puede solaparse totalmente evaluación psicológica en casos reales con metodología científica y/o pasos a realizar para la realización de una investigación.

En un sentido muy similar se expresa Vizcarro (1987) al afirmar que la praxis evaluadora, de ordinario, se realiza en condiciones de presión temporal, lo que se traduce en diferencias importantes entre el proceso de evaluación y el proceso de investigación. Pese a la evidencia de esta disparidad entre lo teórico y lo práctico el proceso empírico de evaluación ha estado más descuidado que el racional o científico investigador, y las investigaciones han incidido más en el cómo debe llevarse a cabo una evaluación que en el cómo se está llevando a cabo.

Es obvio que en una relación personal tal cual se establece en el proceso de evaluación y tratamiento psicológico es imposible el exhaustivo control de variables propio de los diseños de investigación, aunque no es menos cierto que el psicólogo deba procurar ceñirse lo más fielmente a la metodología científica. A pesar del atractivo de mayor rigor metodológico y del poder explicativo del modelo de evaluación conductual, la práctica clínica de este enfoque presenta grandes dificultades; razón por la cual incluso los psicólogos conductistas, en la práctica diaria, usan métodos y técnicas no totalmente acordes con sus presupuestos metodológicos y teóricos (Haynes y O'Brien, 1990; Forns, 1993).

Como se ha visto previamente, muchos autores han situado el proceso de evaluación tradicional como una parte del proceso de evaluación conductual. La descripción y clasificación, finalidades propias de los procesos correlacionales, pueden ser tomados como fase previa a la formulación de hipótesis funcionales, con lo cual habría en el proceso de evaluación conductual, dos momentos de formulación de hipótesis y dos planteamientos en la verificación, que se sitúan en distinto nivel y responden a distintos objetivos.

La discusión que diferenciaba el proceso de evaluación tradicional y el conductista centrándose en la oposición de evaluación de estado versus proceso debe, en opinión de Forns (1993) ser replanteada. En la psicología básica actual las variables de análisis que antaño se habían conceptualizado como estables, se reconceptualizan como modificables. Este es el caso, por ejemplo, de la modificabilidad cognitiva, que conceptualiza la inteligencia como aprendizaje (Calero, 1995c). Este cambio conceptual produce implícitamente un cambio de óptica en la evaluación (Forns y Amador, 1995c).

En la misma dirección que Forns (1993) considero que el proceso de evaluación conductual posee cualidades apreciables de rigor experimental, buena capacidad explicativa-funcional y de control sobre la conducta, mientras que el proceso tradicional es considerado más ambiguo en sus formulaciones, más indirecto en la medición, y dotado de capacidad descriptiva más que explicativa. En cambio, desde un punto de vista aplicado, el proceso de evaluación tradicional, tanto como algunas de sus técnicas, registra un buen índice de frecuencia de uso, aún entre psicólogos no personológicos. En la actualidad, ninguno de los dos diseños del proceso de evaluación satisface, a la par, las exigencias de un enfoque científico de la conducta y las de la aplicabilidad a la realidad clínica, social y educativa.

Si en la década de los 60 se habló de la crisis de la evaluación tradicional, desde los 80 se está hablando de la crisis de la evaluación conductual, para la que en opinión de alguno de sus investigadores (Fernández-Ballesteros y Staats, 1992) urge hallar una respuesta. En esta

dirección, Haynes (1998), ha observado que para muchos pacientes psiquiátricos el diagnóstico resulta insuficiente para el diseño del tratamiento conductual, puesto que existen diferencias importantes entre los distintos pacientes en lo que respecta a los factores causales. Resulta necesaria la valoración conductual previa para asegurar la estrategia de tratamiento más efectiva. Los métodos de evaluación conductual –en comparación con los no conductuales– son menos inferenciales, centrados en factores situacionales, y enfatizan la observación y las habilidades conductuales en el entorno natural. El análisis funcional es la integración de los datos de la evaluación pretratamiento de un paciente, e identifica las relaciones causales importantes y modificables aplicables a conductas específicas de un paciente en concreto. Puesto que llevar a cabo un análisis funcional puede ser complejo, se plantea la utilización del Modelo Funcional Analítico Causal, que sistematiza, integra y cuantifica el análisis funcional. Se propone por tanto este modelo como un método eficaz para organizar las múltiples y complejas variables y relaciones a considerar en un análisis funcional.

En síntesis:

El Psicodiagnóstico de la época anterior a los años 50, se caracteriza por su vinculación a la Psicología Correlacional, imprimiendo este contacto a nuestra materia un interés por tareas y aspectos prácticos, llevados a cabo con la finalidad de describir y clasificar a los sujetos. La posterior aparición del conductismo implicó, entre otros hechos, un cierto cambio en los objetivos del Psicodiagnóstico, debido a que la clasificación resultaba insuficiente, y comienza a requerirse información sobre personas particulares que van a ser tratadas terapéuticamente y sobre los efectos de la terapia aplicada, denominándose a este enfoque Evaluación Conductual. Los nuevos objetivos de la Evaluación Conductual se vinculan a cambios a nivel metodológico, produciéndose un estrechamiento de los lazos con el Método Experimental, lo que ha supuesto modificaciones en el procedimiento a llevar a cabo en el proceso de la Evaluación Psicológica.

Las modificaciones que se han ido produciendo en el objeto de estudio y en la metodología empleada en Evaluación Psicológica, en parte se han producido por las demandas planteadas al psicólogo desde la sociedad, por las propias demandas de los profesionales, así como, debidas a tres causas: 1) Las ocasionadas por el cada vez mayor interés hacia la vertiente orgánica de hombre, apareciendo así la Evaluación Psiconeurológica; 2) Las debidas a la influencia de los intentos interaccionistas, con lo que el Psicodiagnóstico, a parte de ocuparse del individuo (en sus vertientes biológica y psicológica) acaba ocupándose también del entorno social, apareciendo la Evaluación de Ambientes; y 3) Finalmente, se ha derivado de los sencillos diseños y procedimientos para evaluar la eficacia de los tratamientos, a otros más complicados como son los incluidos en lo que hoy en día denominamos Evaluación de Programas de Intervención.

En definitiva, se ha ido produciendo de manera sistemática unos cambios de objetivos e intereses que han supuesto constantes replanteamientos de nuestra disciplina pero, a la vez, han facilitado la apertura de nuevos ámbitos de estudio que hace unos años podrían haber parecido utópicos. Dicho de otra manera, lo que se ha denominado Psicodiagnóstico, y en la actualidad, tras ampliarse el concepto, Evaluación Psicológica se encuentra en constante evolución y se enfrenta a muy diversos tipos de tareas como resultado de los cambios y situaciones que paralelamente van aconteciendo en la Ciencia y en la Sociedad.

5.4. GUIAS PARA EL PROCESO DE EVALUACION (GAP)

En 1995 la Asociación Europea de Evaluación Psicológica (European Association of Psychological Assessment -EAPA-) se propuso elaborar una Guía del Proceso de Evaluación (Guidelines for the Assessment Process: GAP). La comisión encargada de esta labor estuvo compuesta por Fernández-Ballesteros, De Bruyn, Godoy, Hornke, Ter Laak, Vizcarro, Westhoff, Westmeyer y Zaccagnini. El cuarto borrador de las GAP se recoge en la Revista del Colegio Oficial de Psicólogos "Papeles del Psicólogo" en 2003 (en 2001 en la revista "European Journal of Psychological Assessment").

Fernández- Ballesteros, De Bruyn, Godoy, Hornke, Ter Laak, Vicarro, Westhoff, Westmeyer y Zaccagnini (2001, 2003) afirman que por "guía" se entiende *"una sugerencia sobre procedimientos que ayuden a los valoradores y a sus clientes a cumplir los requisitos..."*. Tiene las siguientes ventajas: 1) Ayudar a vencer la frecuente disociación entre conocimiento básico y actividad profesional; y 2) Ayudar a optimizar la práctica profesional, mejorando la valoración y el control de las actividades profesionales y facilitando la formación en evaluación.

El trabajo de esta comisión se ha centrado en detallar las tareas concretas que la persona evaluadora debe llevar a cabo en cada una de las cuatro fases en las que se ha desglosado el proceso de evaluación psicológica:

1. Análisis del caso (evaluación descriptiva)
2. Organización e información de los resultados: elaboración técnica de los resultados e información al cliente o persona
3. Planificación de la intervención
4. Valoración y seguimiento: si se ha intervenido, deben llevarse a cabo nuevos actos de evaluación.

Estas cuatro fases se dividen en distintas subfases donde se detallan los pasos a seguir con las 96 guías, tareas o acciones que la persona evaluadora tiene que realizar durante el proceso de evaluación psicológica (Cuadro 11).

**Cuadro 11. Proceso de evaluación psicológica: estructura básica, pasos a seguir y guías del proceso de evaluación (GAP).
Fernández-Ballesteros, De Bruyn, Godoy, Hornke, Ter Laak, Vizcarro, Westhoff, Westmeyer y Zaccagnini (2001, 2003)**

Comienzo del proceso		
El proceso comienza cuando una persona o institución (cliente) planea a un evaluador que responda a una cuestión o que dé un consejo profesional referente a un sujeto o caso		
Condiciones previas		
El evaluador debe examinar si está cualificado para satisfacer la demanda y si esta se ajusta a criterios profesionales. El evaluador debe observar, además, los principios éticos y los requisitos legales que rijan en el país.		
0. Principios Generales		
1. El evaluador debe asumir la responsabilidad del proceso de evaluación.		
2. El evaluador debe tomar en cuenta los posibles conflictos de intereses que puedan existir entre el sistema de valores del sujeto o cliente y el suyo propio.		
3. La evaluación se lleva a cabo en una situación interpersonal. El evaluador debe tratar al cliente con imparcialidad y con respeto.		
4. El evaluador debe identificar y discutir los asuntos importantes sólo con las personas participantes en el proceso de evaluación.		
5. Durante todo el proceso, el evaluador debe valorar las posibles consecuencias positivas y negativas, así como los efectos colaterales de la evaluación para el sujeto o cliente y para su entorno social.		
6. Al solucionar las demandas planteadas, el evaluador debe seguir, en principio, una aproximación científica.		
7. El proceso de evaluación debe ser lo suficientemente explícito como para que pueda ser reproducido y valorado, así como para que pueda quedar constancia del mismo.		
8. El evaluador debe optimizar la justificación, utilidad y calidad del proceso, así como vigilar las condiciones que puedan distorsionarlo.		
PRIMERA FASE		
1. Análisis del caso (evaluación descriptiva)		
Subíases	Pasos a seguir	Guías del proceso de evaluación (GAP). Tareas o acciones a realizar por el evaluador
1.1. Análisis de las demandas, quejas, metas.	1.1.1. Investigación y evaluación de las demandas del cliente o sujeto.	9. Lograr un adecuado nivel de comprensión de las quejas, demandas y metas del cliente o sujeto.
		10. Considerar el alcance y el nivel de detalle con los que se deben tratar las quejas, demandas y metas del cliente o sujeto.
		11. Averiguar cuál es la percepción del cliente o sujeto sobre la urgencia o severidad del caso.
		12. Explorar tanto los aspectos débiles como los aspectos fuertes del sujeto.
		13. Limitar su exploración a lo que sea pertinente para el caso.
	1.1.2. Síntesis de las demandas del cliente y de los aspectos de la situación	14. Comprobar la suficiencia y precisión de la información básica.
		15. Comprobar si la información recogida resulta suficiente para responder a las demandas planteadas.

	problemática general.		16. Averiguar cómo interactúan la persona y el contexto en el que viven.	
	1.1.3. Acuerdo formal.		17. Informar al cliente o sujeto de las cuestiones que se van a explorar a lo largo de la evaluación.	
			18. Obtener el consentimiento informado del cliente o sujeto sobre las demandas y metas que van a guiar la evaluación.	
1.2. Formulación de hipótesis de evaluación comprobables: conversión de las demandas, quejas, y metas en formulaciones de evaluación.	1.2.1. A partir de la información recogida, formulación de las demandas en términos técnicos.		19. Establecer los constructos psicológicos más idóneos para formular las hipótesis de evaluación sobre el caso.	
			20. Asegurarse de que las hipótesis de evaluación formulada cubren todas las características del problema.	
			21. Distinguir entre hipótesis descriptivas, clasificatorias, explicativas y predictivas.	
			22. Cada una de las hipótesis debe formularse con solidez lógica y teórica, y con referencias empíricas claras.	
			23. Para cada una de las hipótesis formuladas, el evaluador debe generar al menos una hipótesis alternativa rival.	
			24. El evaluador debe jerarquizar las hipótesis de evaluación formuladas de acuerdo con criterios explícitos y apropiados para el caso.	
			25. Revisar los instrumentos de evaluación adecuados para cada constructor y elegir el más apropiado.	
			26. Vigilar que la información a buscar sea pertinente, discriminativa y suficiente con respecto a las hipótesis formuladas.	
			1.2.2. Operativización de las hipótesis mediante instrumentos y procedimientos de evaluación.	27. Tener en cuenta los factores distorsionantes, reactivos y molestos de los procedimientos de evaluación.
		1.3. Recogida de información: recopilación de la información pertinente	1.3.1. Planificación de la administración de los procedimientos de evaluación.	
	29. Preparar la situación de evaluación, así como dar las instrucciones a los sujetos y demás participantes de acuerdo con estándares técnicos y profesionales.			
	30. Aplicar los procedimientos de evaluación de acuerdo con estándares técnicos y profesionales.			
	31. Tener en cuenta los factores que pueden interferir con la adecuada aplicación de los procedimientos de evaluación.			
	32. Comprobar si la recogida de datos se ha efectuado de acuerdo con el plan establecido.			
	33. Comprobar si algún factor ha podido interferir con la apropiada administración de los procedimientos de evaluación.			
	34. Revisar si todos los datos son sólidos y están libres de sesgos y errores de codificación.			
	1.3.2. Aplicación de los procedimientos de evaluación.			35. Revisar los procedimientos de evaluación de acuerdo con estándares técnicos y profesionales.
	1.3.3. Valoración de la aplicación de los procedimientos de evaluación.			36. Comprobar si la aplicación de los procedimientos de evaluación se ha efectuado de acuerdo con el plan establecido.
1.4. Procesamiento de la información: análisis de los datos y toma de decisiones.	1.4.1. Análisis de datos.			
			38. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	
			39. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	
			40. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	
			41. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	
			42. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	
			43. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	
			44. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	
			45. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	
			46. Revisar si los datos son sólidos y están libres de sesgos y errores de codificación.	

información: relación de los datos con las hipótesis planteadas.		35. Valorar la calidad de los datos en relación con las hipótesis planteadas. 36. Analizar e interpretar los datos de las pruebas y demás procedimientos de evaluación de acuerdo con las últimas normas, estándares y conocimientos disponibles.
	1.4.2. Formulación de conclusiones.	37. Comprobar si las conclusiones responden a las hipótesis planteadas. 38. Valorar la información de acuerdo con su importancia para el caso. 39. Justificar la relación existente entre las conclusiones y los datos. 40. Especificar el nivel de confianza que deposita en cada una de las conclusiones.
SEGUNDA FASE		
1. Organización e información de los resultados: elaboración técnica de los resultados e información al cliente o sujeto		
Subfases	Pasos a seguir	Guías del proceso de evaluación (GAP). Tareas o acciones a realizar por el evaluador
2.1. Integración de los resultados: respuesta a las cuestiones planteadas por el cliente o sujeto de la forma más completa posible	2.1.1. Combinación de los resultados en una formulación global del caso.	41. Comprobar cada una de las hipótesis utilizando datos de las fuentes de información apropiadas. 42. Realizar una integración de los datos que resulte adecuada, suficiente y de utilidad para responder a las demandas del cliente. 43. Tener en cuenta la información contradictoria. 44. Formular conclusiones que respondan a las demandas del cliente o sujeto.
	2.1.2. Formulación de conclusiones en relación con las demandas del cliente o sujeto.	
2.2. Informe: presentación oral o escrita del informe	2.2.1. Requisitos para la generación del informe	45. La forma de presentar el informe (oral, escrita o ambas) debe ser la apropiada. 46. El informe debe incluir un resumen con las conclusiones más importantes. 47. El informe debe señalar quien es el autor, el cliente, y el asunto, así como a quién se presentará o enviará. 48. Los datos deben presentarse de acuerdo con las fuentes de información, instrumentos, o procedimientos utilizados. 49. Los datos deben estructurarse siguiendo criterios psicológicos y colocarse en la parte de resultados que habrá en el informe. 50. Los asuntos que sobrepasen las demandas iniciales no deben ignorarse aunque sí considerarse aparte. 51. El informe debe responder a las demandas del cliente con recomendaciones apropiadas.
	2.2.2. Inclusión en el informe de información importante	52. Las demandas del cliente nunca deben quedar sin abordarse. 53. Las fuentes de información, los instrumentos y tests utilizados deben quedar reflejados con el suficiente detalle. 54. El informe debe detallar los pasos y procedimientos utilizados para responder a las

		demandas del cliente. 55. La sección del informe destinada a exponer los resultados debe reflejar el peso dado a la información obtenida y la forma en que se la ha integrado. 56. En la sección del informe sobre los resultados, cada afirmación psicológica que se haga deber estar expresamente basada en los datos recogidos y valorada en función de su importancia con respecto a las demandas del cliente. 57. Las inconsistencias encontradas entre los datos deben discutirse en la sección de resultados.
	2.2.3. Generación de un informe fácil de entender	58. Cada una de las frases empleadas debe ser clara y comprensible para el cliente. 59. El apoyo científico de un enunciado debe explicitarse siempre que exista la posibilidad de que se malinterprete. 60. La terminología técnica utilizada debe clarificarse de la forma apropiada 61. Los elementos descriptivos, comparativos e interpretativos deben distinguirse entre sí. 62. La interpretación de los datos nunca debe dejarse en manos del lector. 63. Las conclusiones deben exponerse con claridad; y las inferencias tentativas, señalarse como tales.
2.3. Discusión y toma de decisiones	2.3.1. Discusión del informe con el cliente, sujeto u otros participantes	64. Discutir las secciones del informe con el cliente o sujeto 65. Tratar con el cliente, sujeto u otras personas participantes las posibles recomendaciones, asegurándose de que las entiende. 66. Reflejar los datos surgidos durante la discusión del informe y de las recomendaciones en la versión final del mismo.
	2.3.2. Analizar si las circunstancias generales aconsejan volver a empezar, finalizar o pasar da intervenir.	67. Si se requiere intervención y el evaluador no está cualificado para llevarla a cabo, el sujeto debe ser enviado al profesional adecuado. 68. Cuando sea pertinente, el evaluador debe enviar un informe a los profesionales oportunos
TERCERA FASE		
3. Planificación de la intervención: si el evaluador considera que es necesario intervenir, hay que llevar a cabo varios actos de evaluación antes de la administración del tratamiento		
Subfases	Pasos a seguir	69. El evaluador debe establecer los criterios y las metas de una intervención provechosa.
3.1. Elección de hipótesis específicas para la intervención	3.1.1. Elección y operativización de la intervención y de las variables para evaluar los resultados.	70. Formular las hipótesis de intervención basándose en los resultados de las evaluaciones previas. 71. En la medida de lo posible, el evaluador debe jerarquizar las hipótesis de intervención siguiendo criterios explícitos y útiles para el caso.

	3.1.2. Revisión y elección de los procedimientos de intervención que mejor se ajustan al caso.	72. Definir operacionalmente las variables relacionadas con las hipótesis de intervención. 73. Inventariar y sopesar los tratamientos posibles y, finalmente, diseñar la intervención. 74. Identificar y evaluar las posibles condiciones facilitadoras e inhibidoras que puedan afectar la intervención. 75. Tratar con el cliente o sujeto las alternativas de intervención y sus posibles consecuencias.
	3.1.3. Elección y evaluación de las variables apropiadas para la supervisión de la intervención	76. Siempre que sea posible, el evaluador debe disponer los medios con los que controlar la marcha de la intervención. 77. Siempre que sea posible, el evaluador debe vigilar la intervención.
REALIZACIÓN DE LA INTERVENCIÓN		
Dado que la intervención no es una actividad específica de evaluación, no se analizará. Aún así, en algunos casos puede que sea necesario realizar una valoración formativa durante la intervención.		
		78. Si es necesario, el evaluador debe vigilar la intervención o tratamiento para evitar desviaciones. Si las desviaciones son inevitables, deben registrarse cuidadosamente.
CUARTA FASE		
4. Valoración y seguimiento: si se ha intervenido, deben llevarse a cabo nuevos actos de evaluación		
		79. El evaluador debe determinar en qué medida se puede valorar la intervención y su puesta en práctica (evaluación valorativa) 80. El evaluador debe informar a los interesados sobre la importancia de realizar una valoración de la misma.
		Gúías del proceso de evaluación (GAP). Tareas o acciones a realizar por el evaluador
Subfases	Pasos a seguir	81. Revisar los datos previos obtenidos en las fases 2 y 3, decidir cuáles son útiles para la valoración y elegir de entre ellos las mejores medidas de resultados.
		82. Comprobar si la intervención se ha llevado a cabo tal como se había planificado. 83. Planificar la valoración con esmero, atendiendo tanto a los datos previamente recogidos ya disponibles, como a los datos adicionales que se deben recoger.
4.1. Recogida de datos sobre los efectos de la intervención.	4.1.1. Inspección de los datos ya existentes.	84. Administrar las medidas elegidas siguiendo lo establecido en 1.3. 85. Recoger, según se establece en 1.3. cualquier otra información sobre cambios positivos o negativos que puedan considerarse resultados de la intervención. 86. Recoger información empleando distintos indicadores de diferentes fuentes.
		87. Comparar los resultados con las demandas del cliente y con las meta de la intervención. 88. Antes de interpretar los efectos de la intervención, el evaluador debe comprobar si dichos efectos realmente pueden atribuir a la intervención. 89. Comprobar si se han producido efectos colaterales. 90. Considerar las posibles explicaciones alternativas de los efectos de la intervención.
4.2. Análisis de los resultados de la intervención	4.2.1. Sobre los datos recogidos y siguiendo lo establecido en 1.4., extracción de conclusiones acerca de los efectos de la intervención.	

4.3. Seguimiento	4.2.2. Informe de los resultados al cliente, sujeto u otros participantes.	91. Tratar los resultados de la valoración con el cliente, sujeto y otras personas participantes. 92. Justificar las conclusiones en relación con las demandas del cliente, las metas de la intervención y los efectos colaterales, de tal forma que los interesados puedan valorarlas. En la medida en que resultan aplicables, ver Guías 2.2.
	4.2.3. Si es necesario, informe escrito al cliente, sujeto o a las otras personas participantes.	
	4.3.1. Planificación del seguimiento de acuerdo con el cliente y el sujetos	93. Planificar el seguimiento, eligiendo objetivos pertinentes y realistas. 94. Discutir y llegar a un acuerdo sobre el plan de seguimiento con el cliente, el sujeto y demás interesados.
	4.3.2. Evaluación del sujeto de acuerdo con el plan establecidos	95. Recoger los datos sobre el sujeto y demás participantes de acuerdo con lo planificado. 96. Recabar información del sujeto y demás participantes sobre posibles consecuencias positivas o negativas previstas.
	4.3.3. Análisis de resultados. 4.3.4. Discusión de los resultados con el cliente, el sujeto y los otros participantes 4.3.5. Si es necesario informar por escrito al cliente, al sujeto, o al resto de los participantes	En la medida en que resultan aplicables, ver Guías 4.2.1. En la medida en que resultan aplicables, ver Guías 4.2.2. En la medida en que resultan aplicables, ver Guías 2.2.1.
FIN DEL PROCESO DE EVALUACIÓN		
El proceso finaliza cuando el evaluador termina su relación profesional, en lo que a tareas de evaluación se refiere, con la persona o institución (cliente) y con el sujeto o caso.		

5.5. CONDICIONANTES QUE INTERVIENEN EN EL PROCESO DE LA EVALUACIÓN PSICOLÓGICA

Existe abundante evidencia empírica sobre el gran cúmulo de variables productoras de sesgos en la investigación que se realiza durante el proceso Psicodiagnóstico-Evaluación Psicológica, tanto de enfoque correlacional como experimental. El análisis de las condicionantes que pueden introducir sesgos durante el proceso, permite establecer tres tipos de agrupaciones según estos procedan del evaluador, del evaluado o de la situación de evaluación.

5.5.1. VARIABLES DEL EVALUADOR

En la evaluación psicológica es difícil controlar los efectos del examinador, convirtiéndolo en una variable experimental más. Sin embargo, resulta necesario establecer un control racional de esta importante fuente de error en el proceso evaluador. Entre las variables dependientes del evaluador se han subrayado las siguientes:

Las expectativas del evaluador: Son de sobra conocidos los efectos que las expectativas del experimentador provocan en la conducta de los sujetos humanos. Los estudios de Rosenthal y Jacobson (1968) pusieron de relieve que las predicciones que los psicólogos realizan en torno al desarrollo cognitivo de los sujetos pueden ser “profecías” que se autocumplan. Los conocimientos que el psicólogo tiene previamente del sujeto, bien directamente, bien transmitidos por el que prescribe la evaluación, y el propio marco teórico del evaluador influyen, al menos hasta cierto punto, en los resultados que se obtengan (Fernández-Ballesteros, 1983d).

El Condicionamiento verbal y Comunicación no-verbal: Otro fenómeno que puede producirse durante el proceso es el condicionamiento verbal del evaluador respecto al evaluado. Como se ha puesto de manifiesto en algunos estudios, refuerzos dispensados por el experimentador en forma contingente a ciertas verbalizaciones del sujeto incrementan éstas. La atención suministrada a ciertos temas, las señales (a veces no verbales) de aprobación o crítica pueden influir en el comportamiento diferencial del sujeto que pasará a actuar en forma reactiva (Fernández-Ballesteros, 1983d).

La presencia o ausencia del evaluador: La presencia o ausencia del evaluador influye en el comportamiento del sujeto, lo cual debe ser tenido en cuenta a la hora de utilizar técnicas de autoaplicación (que se suministran en el domicilio del sujeto). Si bien la presencia del evaluador puede inhibir al sujeto en tareas que exigen la puesta en juego de su creatividad, parece facilitar la respuesta fluida tanto en los autoinformes cuanto en las pruebas de rendimiento intelectual (Fernández-Ballesteros, 1983d).

Grado de especialización y Nivel de entrenamiento del evaluador: Un tema que ha despertado cierto interés y polémica es el relativo a las competencias profesionales durante el proceso. Al hilo de esto, se circunscribe una polémica entre los distintos autores, algunos de los cuales consideran que el mismo profesional debe realizar la evaluación y el tratamiento, mientras que otros se posicionan por la especialización. Pelechano (1988), efectuando un

parangón con la Medicina, deja entrever la deseabilidad de una mayor especialización. Distingue el autor entre el clínico (profesional que se ocuparía de todas y cada una de las fases del proceso) y el psicólogo evaluador, ocupado fundamentalmente en la etapa de medición de ciertos parámetros. Ello implica una clara crítica a los modelos en uso que abogan por una evaluación integrada que se prolonga en fases de intervenciones y tratamientos. El problema no está resuelto por el momento, ya que a mi juicio ambas posturas tienen sus ventajas y limitaciones. Mientras que la especialización puede estimular que se pierda de vista la globalidad del sujeto, permite, sin embargo, una mejor profundización en aspectos diversos del proceso de evaluación, imposibles de conocer y manejar por un sólo profesional. Y esto aún será más evidente en la medida que se complejicen los métodos y las técnicas evaluativas. Incluso en algunos casos, como puede ser el de la evaluación de intervenciones, la derivación a otro profesional permitiría evitar errores de sesgo e implicación personal. Como parece lógico pensar el entrenamiento del evaluador es una condición importante en el transcurso del proceso. Sin embargo, sus efectos en un sentido u otro no parecen claros, ya que mientras algunos sostienen que un buen entrenamiento neutraliza toda fuente de sesgos, otros mantienen que éste no parece mejorar las predicciones derivadas de la evaluación (Fernández-Ballesteros, 1983d). Un estudio que ha analizado el papel del entrenamiento del entrevistador y su dominio de los instrumentos de evaluación que emplea, subraya el papel positivo de este dominio que implica necesariamente el seguimiento de las normas estandarizadas (Bartholomew, 1992).

Características físicas y psicológicas del evaluador: De entre las características psicológicas se ha destacado que la *actitud de vigilancia* tiene una incidencia negativa en los resultados, ya que parece coartar el comportamiento durante la exploración (Fernández-Ballesteros, 1983d). Así mismo, se ha llamado la atención sobre las variables de *personalidad* del psicólogo que, bien a través de su conducta interpersonal o su proceso de juicio decisional, determinan las hipótesis diagnósticas. Además los *prejuicios* culturales, sociales, económicos, étnicos... pueden incidir sobre la lectura que haga el examinador (Avila Espada, 1992b). Un modo de instrumentalizar, en parte, el control de estas variables es que el evaluador tenga conocimiento de las mismas, como sugieren diversos estudios. Por lo que se refiere a las *características físicas*, las más relevantes parecen ser el *sexo* y la *apariencia externa*. Se ha demostrado que examinadores del sexo contrario al del sujeto en evaluación, inhiben el relato de determinados temas, así como que modos de vestir convencionales o desenfadados influyen en las respuestas que dan los sujetos (Fernández-Ballesteros, 1983d). Estudios empíricos recientes sobre las experiencias del evaluador y del evaluado en el proceso de evaluación, llevados a cabo con una correcta metodología cualitativa, permiten concluir que ambos están implicados en un proceso continuo y circular de feed-back que compromete las decisiones diagnósticas y terapéuticas (Ward, 2008).

5.5.2. VARIABLES DEL EVALUADO

Otras experiencias previas de evaluación: Otra variable que puede influir sobre la situación de evaluación actual es la existencia de experiencias previas de evaluación por parte del sujeto. Es importante la exploración de este aspecto ya que si estas experiencias previas han sido negativas podrán condicionar negativamente el comportamiento del sujeto y sus respuestas en

los instrumentos de evaluación administrados. Debido a ello, resulta conveniente explorar este aspecto en los momentos iniciales del proceso de evaluación psicológica.

Motivación para la exploración: Para Fernández-Ballesteros (1983d) la condición más importante es, sin duda, el estado motivacional del sujeto ante la exploración. El hecho de que éste acuda voluntariamente o no, en demanda de orientación o tratamiento es ya un primer dato que es necesario tener en cuenta. La evidencia empírica apoya el supuesto de que los sujetos que asisten voluntariamente al psicólogo suelen mostrarse más responsivos y colaboradores que aquellos que lo hacen involuntariamente e incluso, estos últimos pueden presentar actitudes contrarias al examen e incluso falsificar sus respuestas al ser testados. Hay que tener en cuenta que los humanos son perfectamente capaces de falsear su comportamiento, incluso ante técnicas que se consideran enmascaradas y no voluntarias. Debido a todo ello, es importante que el psicólogo se plantee antes de la evaluación, el estado motivacional del sujeto frente a la exploración, no comenzando ésta hasta cerciorarse de que la participación va a ser la adecuada, lo cual en algunos casos implicará una preparación previa del paciente en orden a incrementar su nivel motivacional.

El lenguaje: Otro elemento central en la práctica de la evaluación psicológica está relacionado con el lenguaje, no sólo en lo que se refiere a los elementos implícitos de comunicación, sino también los elementos paralingüísticos, los condicionamientos verbales, las constricciones del entorno, las influencias de los estados emocionales y aptitudes lingüísticas (Matud, 1993b). El análisis del lenguaje es un aspecto esencial, ya que como indica Pelechano (1988c) el lenguaje se ha hecho tanto para decir verdades como para contar mentiras, siendo una función del evaluador discriminarlas.

Deseabilidad Social: Los dos condicionantes previos hacen referencia a la posibilidad de falseamiento voluntario del evaluado con la finalidad de distorsionar deliberadamente su imagen ante el psicólogo. Sin embargo, en ocasiones el sujeto no actúa deliberadamente cuando responde dando una imagen positiva de sí mismo. A este hecho suele denominarse deseabilidad social. La deseabilidad social es la tendencia del sujeto a responder de forma socialmente aceptable, afín de transmitir una imagen positiva de sí. Este sesgo suele darse especialmente en los autoinformes, es decir, cuando se solicita al sujeto que informe sobre sí mismo. Un modo de neutralizar este factor puede ser la aplicación de múltiples técnicas de indagación, es decir, empleando distintos tipos de instrumentos de evaluación para medir la misma variable.

Características psicológicas y físicas del evaluado: Características psicológicas del sujeto a evaluar, como la extraversión o el neuroticismo, parecen también influir en cómo se comporta el sujeto durante la exploración. No obstante, hay que tener en cuenta que la Evaluación se dirige también a la exploración de estos comportamientos diferenciales por lo que deben ser considerados como datos de evaluación y no tanto como sesgos del proceso. Por otro lado, las características físicas como el sexo, y, sobre todo, la edad, intervienen de manera notable durante el proceso. Las influencias negativas que el sexo del sujeto, por oposición al del evaluador, ejerce sobre sus respuestas, deberá ser controlado. La edad es otra variable relevante, por lo que es importante adaptar el contexto y las técnicas de aplicación a la edad del

sujeto, de tal modo que en las distintas edades los sujetos puedan lograr una buena adaptación a la situación de evaluación (Fernández-Ballesteros, 1983d).

5.5.3. VARIABLES DE LA SITUACIÓN

La situación de evaluación implica múltiples variables procedentes tanto del ambiente físico en el que ésta se realiza, de las técnicas que son aplicadas, así como de contextos más amplios al de la situación de evaluación.

Medio físico de la evaluación: El medio físico en el que se realiza la evaluación debe contar con espacio, luminosidad, temperatura, nivel de ruido, etc. adecuados a la tarea que el psicólogo y el sujeto deben realizar. Existe evidencia de que condiciones físicas adversas inhiben el comportamiento tanto del evaluador como del sujeto evaluado (Fernández-Ballesteros, 1983d), lo que distorsionará la validez del proceso.

Tipo de técnicas aplicadas: Otro factor de influencia es el *tipo de material* con el que se registra la información sobre el sujeto, ya que los distintos tipos de técnicas influyen diferencialmente en las respuestas que da el sujeto. Para Fernández-Ballesteros (1983d), el tipo de consigna, el tiempo de aplicación, la vulnerabilidad de las pruebas, su enmascaramiento, etc. son características influyentes en los resultados de la exploración por ser, precisamente, el material de evaluación.

El contexto institucional donde se realiza la evaluación: El contexto específico donde se realiza la evaluación dota al proceso de funcionalidad, finalidad y significaciones peculiares (por ejemplo, los diferentes contextos institucionales, docentes o privados en los que se realizan las prácticas evaluativas, los significados micro y macrosociales de dicho proceso evaluativo, etc.) (Avila-Espada, 1992b).

5.5.4. LA CALIDAD DEL PROCESO DE EVALUACIÓN PSICOLÓGICA

5.5.4.1. La eficacia de los evaluadores en tareas de clasificación, predicción, prescripción de tratamiento...

La investigación sobre la calidad del proceso de evaluación psicológica arroja diversos resultados (Moreno, 2003):

1. En cuanto a la *eficacia de los evaluadores en tareas de clasificación* (Vizcarro, 1987; Silva, 1988; Godoy, 1996; Godoy y Gavino, 1999):
 - La seguridad de los evaluadores al emitir los “juicios diagnósticos” no concuerda con la exactitud de los mismos.
 - No existen diferencias respecto a las variables de personalidad de los evaluadores, ni la edad, años de experiencia, estar o no psicoanalizados.
 - Las tareas a realizar o las características de los sujetos sobre los que se realiza la clasificación no tienen relación con la habilidad del buen profesional para clasificar.

- Sí hay diferencias con respecto a la calidad de los instrumentos utilizados, pero por debajo de lo esperado.
 - No existe relación con la cantidad de información que el evaluador recibe para resolver el caso, ni con el tipo de formato en que se piden las conclusiones.
2. En relación a las *tareas predictivas* realizadas por los psicólogos, los resultados no son mejores que los anteriores, ya que para poder predecir es necesario clasificar.
 3. En cuanto al estudio sobre el papel de los clínicos al *prescribir tratamientos*, las conclusiones a las que se llegan son parecidas. El diagnóstico indica cuál es el tratamiento adecuado.
 4. Referido al estudio sobre la *bondad del proceso de evaluación conductual*, uno de los problemas característicos ha sido la concordancia entre jueces en relación a la identificación de conducta-problema de los pacientes.

A pesar de todo, existen buenos profesionales, que presentan diferencias individuales y son capaces de distinguir su vida, creencias y valores y expectativas de aquello que es mejor para su paciente; en consecuencia, el rigor en las tareas durante el proceso de evaluación psicológica, no sólo va a depender de factores atribuibles a los modelos, sino también a los clínicos o psicólogos que las instrumentan y a los contextos de intervención para los que se realiza la tarea evaluativa.

5.5.4.2. La relevancia de la capacidad de metacognición del evaluador

Una línea de investigación reciente se ha centrado en la importancia de las cogniciones y el conocimiento del evaluador para la formulación del diagnóstico y la toma de decisiones, así como la relevancia de la capacidad de evaluación de la metacognición del evaluador, en la calidad del proceso de la evaluación psicológica.

Las tareas a realizar en función de los objetivos específicos de cada una de las fases del proceso de evaluación incluyen las cogniciones del evaluador, el conocimiento desplegado por él que aplica en múltiples momentos en cada caso, siguiendo el método inductivo-hipotético-deductivo. Así, en la fase de formulación de hipótesis y deducción de enunciados verificables, el clínico pasa por un momento inductivo, no exento de creatividad, al que seguirá un momento deductivo en el que se planteará deducciones verificables sobre el caso, etcétera (Gabucio y Lichtenstein, 2005).

Así pues, es importante prestar atención al estudio del sistema complejo de procesos cognitivos interactuantes en la mente del evaluador clínico. Éste genera activamente la información, la codifica, la transforma y manipula de múltiples formas (no sólo con razonamiento), a fin de avanzar en el conocimiento pertinente a la tarea evaluadora, y de esa manera resolver los difíciles problemas que ésta acarrea (Salonen, Vauras y Efklides, 2005).

La evaluación clínica es una actividad genuinamente cognitiva que requiere ejercitar el razonamiento silogístico, lo que exige planificar, percibir, recordar, pensar, analizar, sintetizar, inferir, valorar, enjuiciar, tomar decisiones, etc. El ser humano – con mayor razón si cabe el profesional clínico – no se limita a ejercer un conjunto de cogniciones sin relación entre sí, sino que trabaja desde un sistema cognitivo organizado de modo muy complejo, cuyos elementos interactúan entre sí dinámicamente (Flavell, 1984). Estas actividades cognitivas del evaluador,

subyacentes a sus funciones como profesional de la clínica, deberían poder ser revisadas, analizadas, perfeccionadas. Estas actividades actúan organizadamente originando procesos metacognitivos del propio saber, y han sido Cooper, Grocutt, Deepack y Bailey (2007) los que los han estudiado en el ámbito de la psicología clínica.

Los conocimientos (knowledges) del profesional y sus cogniciones (cognitions) en la práctica profesional son inseparables, demandan nuestra atención y deberían ambos ser considerados con atención para su adecuado análisis y valoración. La atención del evaluador clínico debiera ir más allá de los resultados y lograr explorar los procesos implicados, es decir, las cogniciones efectuadas por el clínico en el desarrollo de su funciones, saber cuáles son los pasos dados adecuadamente y cuáles los inadecuados o mejorables.

Como telón de fondo de las competencias del clínico evaluables, además de sus conocimientos generales y específicos ligados al caso, es importante considerar un conjunto de capacidades imprescindibles en él. Entre ellas, se puede mencionar, su habilidad práctica, capacidad crítica y analítica, capacidad de solución de problemas y creatividad, capacidad de comunicación, habilidades sensibilidad hacia los valores éticos, habilidades de selección, organización y síntesis de las informaciones, capacidad interpretativa y de utilización de los datos cuantitativos y cualitativos, además de su capacidad de observación científica, autonomía en la toma de decisiones, capacidad de trabajo en equipo y capacidad de autocrítica y autoevaluación.

Son estos diversos y complejos procesos cognitivos que forman parte de la metacognición los que han llevado a replantearse en el momento actual la vinculación entre metacognición y nivel sociopersonal del evaluador, así como metacognición y afectos (Efklides, 2008; Koriat, 2007). Sin duda, existe una vía incipientemente explorada de investigación que relaciona el carácter dinámico de la auto-regulación y co-regulación de las cogniciones, emociones y conductas del profesional evaluador (Efklides, Kourkoulou, Mitsiou y Ziliaskopoulou, 2006; Koriat, 2007)

El término metacognición fue introducido por Flavell (1976) en la psicología, y desde entonces hasta ahora los desarrollos en este campo han acumulado abundante investigación (Efklides, 2008). La metacognición es el conocimiento del conocimiento, atañe a las dimensiones cognitivas que conforman el proceso de avanzar en el conocimiento (Aleman y Laroí, 2005). Esta capacidad especial que está en el núcleo de la autoreflexividad, el autoconocimiento y la autocrítica, no es fácil de ejercitar, y en ella se funda el metaconocimiento de las evaluaciones efectuadas, es decir, la metaevaluación (Efklides, 2006; 2009).

La metacognición se sustenta sobre la capacidad humana de analizar los propios procesos mentales, evaluándolos para después modificarlos. Existen en uno mismo procesos cognitivos generales presentes en cualquier situación de resolución de problemas de los cuales un sujeto toma conciencia y autovalora. En cualquier caso, el funcionamiento metacognitivo depende de la accesibilidad del pensamiento y de la utilización flexible de las propias habilidades de pensamiento (Bandura, 1987; Efklides, 2006, 2009). Y todo ello aun constatando que las personas en gran parte son inconscientes de sus procesos cognitivos mayormente automáticos.

La metacognición es la habilidad para controlar los procesos de conocimiento, observarlos, revisarlos y modificarlos (Khun, 2000). La metacognición es tarea difícil, cuesta trabajo ejercitarla, sobre todo si no se tiene bien claro para qué puede servir concienciar aquello

que con frecuencia se realiza de manera espontánea, tal vez demasiado rápida, y sobre todo automática. Es más accesible a la conciencia saber qué se hizo y cómo, que saber en qué razonamientos se encontró apoyo no sólo para hacerlo sino también para decidir hacerlo y las inferencias efectuadas. Sin embargo, dichas cogniciones del evaluador están estrechamente relacionadas no sólo con el producto de la evaluación, sino con la mayor o menor calidad científica y sanitaria de los resultados. Como enfatizan Saldaña y Aguilera (2003), sin una adecuada planificación, supervisión y evaluación de la tarea, es difícil que el sujeto desarrolle o utilice metaconocimiento alguno.

Tres dimensiones se distinguen en torno a los procesos metacognitivos: a) el conocimiento metacognitivo, que ofrece datos sobre distintos aspectos de la cognición (la tarea, la persona, las estrategias; b) las habilidades metacognitivas, en particular las de autorregulación, como son planificar, supervisar, revisar; y c) la experiencia metacognitiva o acceso consciente reflexivo ejercido sobre las propias cogniciones y actuaciones. La personalidad del clínico es una de las variables del sujeto que permite definir el metaconocimiento funcionalmente. Otras son: su nivel de conocimientos, las creencias, el nivel de habilidades, los estilos atribucionales, las motivaciones, etc. (Khaneman, 2003). Estas dimensiones han sido principalmente investigadas por Efklides (2001, 2002a; 2002b; 2006; 2008; 2009) y Koriat (2007), aportando resultados incipientes a la interrelación de las mismas. Sin duda, que en un futuro próximo nuevos desarrollos vincularán cada una de ellas a las propuestas conceptuales de la teoría de la mente (TOM), la metacognición de los conocimientos (MK), las experiencias metacognitivas (ME) y las habilidades metacognitivas (MS) (Efklides, 2008, Koriat, 2007).

La eficacia de las cogniciones de los evaluadores clínicos en tareas de clasificación diagnóstica no es la deseable. Se han hallado importantes deficiencias y errores al respecto. Entre ellas, como expresan algunos autores (Kahneman, 2003; Godoy, 1996): escasa concordancia entre la seguridad al emitir los juicios y la exactitud de los mismos; las variables cognitivas o de personalidad de los jueces no marcan diferencias significativas en la calidad resultante de los juicios diagnósticos; no se halla relación entre las características del paciente y la habilidad del clínico para juzgarlo en el diagnóstico; y ni siquiera es significativa la cantidad de información barajada al respecto por el clínico (Salonen, Vauras y Efklides, 2005).

Varios sesgos han sido estudiados en el proceso de atribuir una categoría diagnóstica a un sujeto. Entre ellos, además del sesgo de la perspectiva actor-observador, también llamado de correspondencia, en la configuración de las explicaciones suelen surgir otros sesgos atribucionales, como son (Worchel, Cooper, Goethals y Olson, 2003): 1) Inferencias espontáneas de rasgos (IER), que se produce cuando observamos un comportamiento y hacemos automáticamente inferencias de rasgos. Como quiera que estas inferencias son automáticas, ocurren fuera de la conciencia no cabe aplicar el esfuerzo por controlarlas; y 2) Sesgo de preponderancia (o "saliencia"), por el cual cualquier estímulo vívido, destacado, en la práctica goza de mayores probabilidades de recibir la atribución causal del comportamiento observado.

Dada la falibilidad del juicio humano, sería deseable aceptar la premisa de que es probable equivocarse, sobre todo a la hora de tomar decisiones importantes como las que incumben a los clínicos, que a veces yerran por exceso de confianza en sus juicios y carencia de autocritica (Limón, 2005). Una excelente revisión sobre los errores, falsas creencias y justificación infundada de la toma de decisiones ha sido realizada por Tavis y Aronson en 2007.

Las habilidades metacognitivas son macroestrategias vinculadas a las estrategias de organización de la información, y a las estrategias de regulación o control cognitivo y afectivo-emocionales (preferencias, estilo personal, variables disposicionales...) activadas en el momento de ejercitar el metaconocimiento (Efklides, 2009). Las habilidades metacognitivas se refieren al conjunto de estrategias a fin de controlar la cognición, lo que sin duda implica (Shinamura, 2000), estrategias de orientación, estrategias de regulación cognitiva, monitorización para la ejecución de un plan de acción, y estrategias para la evaluación de la propia tarea (Efklides, 2008; Veenman y Elshout, 1999).

La metacognición del evaluador también debe ser objeto de evaluación. No obstante dicha evaluación de la metacognición es una actividad no exenta de dificultades (Pintrich, Wolters y Baxter, 2000; Schraw, 2000). Se precisan instrumentos de evaluación metacognitiva que en los últimos años están empezando a diseñarse (Troyer y Rich, 2002). Autores como Sandi-Urena (2009) han diseñado y validado uno de los instrumentos de evaluación sobre metacogniciones e intervenciones metacognitivas, concluyendo que es posible determinar algunas habilidades cognitivas que permiten tomar conciencia del proceso de la cognición en el proceso de la evaluación psicológica.

Se pueden destacar tres métodos de mejora de las habilidades metacognitivas: el modelamiento metacognitivo, el análisis y discusión metacognitivos, y la autointerrogación metacognitiva. La optimización de las habilidades metacognitivas del clínico se consigue implementando, a poder ser concatenadamente, las estrategias siguientes: 1) Reflexión sobre su propia actuación profesional y los mecanismos cognitivos ejercitados en ella; 2) Discusión y debate grupal tendente a identificar las metacogniciones, analizarlas y discutir su adecuación o inadecuación en cada caso; 3) Mantenimiento de la tensión de búsqueda del desarrollo de las cogniciones más rigurosas y eficaces; y 4) Autoaplicación permanente de aquellas mejoras halladas en habilidades y estrategias a fin de asegurar su optimización progresiva (Efklides, 2009).

En síntesis, de los estudios llevados a cabo en los últimos años sobre la importancia de las cogniciones y el conocimiento del evaluador en el proceso diagnóstico y la relevancia de la evaluación de la metacognición del evaluador, se puede concluir que es necesario que el clínico, en su proceso de atribución diagnóstica, conozca sus propias capacidades cognitivas y metacognitivas, así como las estrategias y macroestrategias realizadas en dicho proceso.

5.6. RESUMEN: FASES, OBJETIVOS, TAREAS Y TÉCNICAS EN EL PROCESO DE LA EVALUACIÓN PSICOLÓGICA

Con la finalidad de clarificar las fases, objetivos, tareas y técnicas del proceso de evaluación psicológica, sintetizando lo dicho hasta el momento, presentaré a continuación un esquema de estas fases, que tiende a integrar la perspectiva correlacional y experimental del proceso, metodologías por las cuales se optarán, o complementarán en función de los objetivos de evaluación (clasificación, selección, tratamiento o cambio...). Esta perspectiva integradora de enfoques por la que optamos cuando nos situamos en la praxis evaluadora, la adoptaremos así mismo en el transcurso de la materia (selección de instrumentos de diversos tipos, elección de técnicas de tratamiento en función de la problemática y la eficacia de los tratamientos evidenciados en distintos enfoques para esa problemática...) (Cuadro 12).

Cuadro 12. Fases y tareas en el proceso de Evaluación Psicológica

FASE DE RECOGIDA DE INFORMACIÓN	
Contacto inicial	Objetivo: Clarificar el Motivo de consulta, la historia del sujeto y la situación actual. Técnicas: Entrevistas al sujeto, allegados, otros profesionales... Historia Clínica /Evolutivo-Académica-Profesional. Observación.
Hipótesis diagnóstica	Elaboración de una Teoría sobre el caso (en función del marco teórico). Hipótesis: Cuantificación, Semejanza y Asociación predictiva. Factores Etiológicos y Actuales.
Diseño de la evaluación	Selección de las Variables objeto de estudio (en función de las hipótesis) Tipos de Variables: del individuo, del evaluador, del contexto familiar-escolar-sociocultural, de las tareas o conflictos por resolver: Por ejemplo, inteligencia, pensamientos, creencias, rasgos de Personalidad, aptitudes, creatividad, ambiente, habilidades sociales, condiciones biológicas (psicofisiológicas / neuropsicológicas) ... Elección de los tests y las técnicas adecuadas de evaluación (en función del marco teórico) Tipos de Técnicas: Observación, Autoinformes, Proyectivas, Objetivas, Subjetivas. Empleo de técnicas: Globales y Focales. Planificación de la batería de pruebas y programación de su aplicación.
Aplicación de las técnicas	Entrevistas de administración de tests y otras técnicas de evaluación seleccionadas para evaluar las variables objeto de estudio.
FASE DE PROCESAMIENTO DE LA INFORMACIÓN	
Corrección e interpretación de datos	Análisis de los datos de la/s Entrevista/s, y Corrección-Interpretación de tests-técnicas aplicadas que conduce a la Integración Diagnóstica (si se requiere más información se aplicarían nuevas técnicas de recogida de datos). Los Resultados se contrastan con las hipótesis iniciales.
Integración diagnóstica y orientación	Formulación de un Diagnóstico: Descripción, Clasificación, Predicción Integración: Dinámico y Conductual (Inferencial y Descriptivo): • Explica desde la historia del individuo cómo ha llegado a ese punto • Explica las condiciones que actualmente mantienen la/s conducta/s problema Clarificación de la Orientación: Toma de Decisiones: Por ejemplo, reeducación instrumental, programa de habilidades sociales, tratamiento psicoterapéutico...

Elaboración de un informe	• Datos de identificación personal. • Motivo de consulta y otros problemas que aparecen. • Áreas de funcionamiento positivo. • Exploración realizada (variables medidas y técnicas empleadas). • Comportamiento del sujeto en el proceso de la evaluación. • Resultados obtenidos con los instrumentos aplicados. • Diagnóstico (conclusiones). • Orientación-Recomendaciones.

FASE DE DEVOLUCIÓN DE LA INFORMACIÓN	
Modalidad	Oral: Entrevista de Devolución / Escrita: Informe Psicológico Esquema: Motivo de consulta (objetivos de la evaluación). Exploración realizada. Resultados obtenidos. Diagnóstico y orientación-recomendaciones. Se han escrito variadas obras sobre la devolución y el informe (Avila-Espada, 1992b; Caballo, 2006; Fernández-Ballesteros, 1992h; Forns, 1993; Frank de Verthelyi, 1989; Lichtenberger, Mather, Kaufman, y Kaufman, 2006; Maganto 1990, 1995b; Maloney y Ward, 1976; Martorell, 1985, 1988e; Pelechano, 1988; Rodríguez, 2003), a las que se sugiere acudir para ampliar información sobre este tema.
Variaciones	• En función de las metas de evaluación (diagnóstico, orientación, selección o tratamiento). • En función del que recibe la información (sujeto, padres, maestro, institución, otro profesional...)

FASE DE DISEÑO Y APLICACIÓN DEL PROGRAMA DE INTERVENCIÓN	
Diseño del plan de intervención	• Determinación de los objetivos terapéuticos o de cambio (en función de las conclusiones diagnósticas). • Elección de las estrategias de intervención o tratamiento. • Temporalización de la aplicación del plan de intervención. • Selección de instrumentos de evaluación continua del tratamiento.
Aplicación del tratamiento o programa	• Puesta en práctica del plan de intervención o tratamiento. • Evaluación continua en distintos momentos de la aplicación del tratamiento o del programa de intervención.

FASE DE EVALUACIÓN DE LA INTERVENCIÓN	
Valoración de resultados	Nueva evaluación con los mismos (u otros) instrumentos que en la evaluación pretratamiento para medir los efectos del tratamiento en las variables objeto de intervención
Seguimiento	Temporalización de nuevas evaluaciones a 6 o 12 meses de la finalización del tratamiento.

Como he comentado previamente, en todo este proceso pueden influir muchos factores condicionantes que pueden afectar a la validez y fiabilidad de la información recogida, por lo que a continuación esquematizamos los sesgos o fuentes de distorsión que se pueden derivar de la figura del evaluador, del evaluado y de la situación de evaluación (Cuadro 13).

Cuadro 13. Condicionantes que pueden afectar en el proceso de evaluación

CONDICIONANTES QUE PUEDEN AFECTAR EN EL PROCESO DE EVALUACIÓN	
Evaluador	• Las expectativas del evaluador (Efecto Rosenthal). • El condicionamiento verbal y la comunicación no verbal. • La presencia o ausencia del evaluador. • El grado de especialización y nivel de entrenamiento. • Características psicológicas y físicas.
Evaluado	• Otras experiencias previas de evaluación psicológica positivas o negativas. • Motivación para la exploración. • El lenguaje, por su posible utilización para ocultar información. • La deseabilidad social. • Características psicológicas y físicas.
Situación	• El medio físico donde se opera la evaluación (espacio, temperatura, luz, intimidad...). • El tipo de técnicas aplicadas. • El contexto institucional donde se realiza la evaluación.

CAPÍTULO 6.

GARANTÍAS CIENTÍFICAS Y ÉTICAS DE LA EVALUACIÓN PSICOLÓGICA

- 6.1. Introducción
- 6.2. Garantías científicas de la evaluación psicológica
 - 6.2.1. Test, técnicas y otros procedimientos de evaluación
 - 6.2.2. Procedimiento para la elaboración de un test
 - 6.2.3. Interpretación de las puntuaciones utilizadas
 - 6.2.4. Propiedades de los tests y de las técnicas estándar de medida
 - 6.2.4.1. La fiabilidad
 - 6.2.4.1.1. Fiabilidad test-retest o estabilidad temporal del test
 - 6.2.4.1.2. Consistencia interna
 - 6.2.4.1.3. Fiabilidad del evaluador o interjueces
 - 6.2.4.2. La validez
 - 6.2.4.2.1. Validez de contenido
 - 6.2.4.2.2. Validez criterial
 - 6.2.4.2.3. Validez de constructo
 - 6.2.4.2.4. Últimos desarrollos del concepto de validez: Estándares del APA (1999)
 - 6.2.5. Teoría de la Generalizabilidad TG
 - 6.2.6. Aplicaciones de la Teoría de Respuesta al Ítem TRI
- 6.3. Garantías éticas en la evaluación psicológica
 - 6.3.1. El código deontológico del psicólogo
 - 6.3.2. Directrices asociadas al uso de los tests
 - 6.3.2.1. Directrices internacionales para el uso de los tests
 - 6.3.2.2. Directrices para la traducción y adaptación de tests
 - 6.3.2.3. Directrices internacionales para la evaluación informatizada y test aplicados por internet.
- 6.4. Resumen: Garantías científicas y éticas en la evaluación psicológica

6.1. INTRODUCCIÓN

En este capítulo, en primer lugar, se abordan una serie de cuestiones vinculadas con las garantías científicas requeridas en un proceso de evaluación psicológica. En este contexto se especifican aspectos referentes al modo de interpretación de las puntuaciones directas, criterio y norma, así como aspectos relacionados con la fiabilidad, la validez de los instrumentos. En segundo lugar, se tratan una serie de aspectos éticos de primordial importancia a tener en cuenta cuando se realiza una evaluación psicológica.

6.2. GARANTÍAS CIENTÍFICAS DE LA EVALUACIÓN PSICOLÓGICA

6.2.1. TEST, TÉCNICAS Y OTROS PROCEDIMIENTOS DE EVALUACIÓN

No todos los procedimientos de evaluación psicológica son tests. Existen otros dispositivos de recogida de información que vamos a denominar técnicas o instrumentos de evaluación (Fernández-Ballesteros y Calero, 2004). La palabra test etimológicamente procede de la palabra latina *testis*, que significa prueba o ensayo (Del Rincón, Arnal, Latorre y Sans, 1995), siendo en su procedencia un concepto muy genérico. Por ello, diversos autores como Pelechano nos brindan unas matizaciones que son necesarias para clarificar estos conceptos. *Instrumentos y técnicas* serían vocablos equiparables que comparten la característica común y diferencial con respecto a los tests de que son *procedimientos no cuantificados ni tipificados* para efectuar un diagnóstico psicológico. El *test*, por el contrario, es un *instrumento sistemático y tipificado que compara la conducta de dos o más personas* (Pelechano, 1976).

Un test, como señala Cronbach (1990), es un procedimiento sistemático para observar la conducta y describirla con la ayuda de escalas numéricas o categorías establecidas... El evaluador recoge información preguntando y observando a todas las personas de la misma manera, en la misma o en comparables situaciones. Esta definición engloba los cuestionarios, a través de los cuales se obtienen informes sobre la personalidad, procedimientos para observar la conducta social, aparatos de medida de la coordinación o, incluso, registros sobre productos. Dos extremos conviene que sean resaltados: por un lado, el requisito de que el test presente escalas numéricas o categorías establecidas; por otro lado, que a todas las personas se les debe preguntar u observar de la misma manera.

Recientemente, Urbina (2007) define el test como un procedimiento sistemático para obtener muestras de conducta significativas del funcionamiento cognitivo o afectivo de la persona y para puntuar y evaluar estas muestras de acuerdo con normas tipificadas. En definitiva, *los tests son procedimientos de medida cuyo material, forma de administración, puntuación, corrección y valoración son estándar y de los que se derivan puntuaciones normativas*. Así, las técnicas o instrumentos se diferencian de los tests por no cumplir alguna de las características mencionadas (Fernández-Ballesteros y Calero, 2004).

6.2.2. PROCEDIMIENTO PARA LA ELABORACIÓN DE UN TEST

El primer factor importante relacionado con las garantías científicas de un instrumento hace referencia al procedimiento seguido para su construcción. La elaboración de un instrumento de medida como un test es un proceso complejo en el que intervienen distintos

factores. Para garantizar la adecuabilidad de un test, Del Rincón et al. (1995) sugieren seguir varios pasos en el proceso de construcción que se sintetizan en el cuadro 14.

Cuadro 14. Proceso para la elaboración de un test I (Del Rincón et al. 1995)

PROCESO PARA LA ELABORACIÓN DE UN TEST	
Elección y definición del constructo o rasgo a medir	Elegido el rasgo a medir es necesario definirlo conceptualmente y operativizarlo, ya que el rasgo no suele ser directamente observable. Esta fase ha de dar respuesta a preguntas tales como ¿Qué se va a medir? ¿Qué es el rasgo?
Planificación del test	Es la fase en la que el investigador toma decisiones como las siguientes: tipo de test, medio material, tipo de respuesta, tiempo de aplicación y longitud del test.
Fuentes de información	Las fuentes de documentación sobre el rasgo ayudan a perfilar la concepción general del instrumento y a evitar errores.
Redacción de los elementos o ítems del test	Se redactan ítems adecuados para explorar las conductas que pueden aportar información sobre el rasgo a medir. Resulta aconsejable redactar un número amplio de ítems, tres veces superior al número de ítems que tendrá el test definitivo.
Elección de los ítems	En esta fase varios expertos en la materia seleccionan los ítems que configurarán la redacción inicial del tests. Previamente se les informa sobre el objetivo de la medida, la definición del constructo o rasgo a medir, la finalidad del test y las características del grupo al que se aplicará. Esta primera selección de ítems configura una primera versión del test y se aplica a una pequeña muestra de sujetos.
Edición experimental del test	Teniendo en cuenta los resultados de la fase anterior, se estructura la edición experimental del test y se aplica a grupos de sujetos que posean las características definidas y también a grupos con niveles superiores e inferiores a los establecidos en los prerrequisitos, con el objeto de valorar si cada ítem es adecuado para el nivel que se elaboró. Posteriormente se realiza un análisis de las características técnicas de cada ítem. Para realizar la selección hay que tener en cuenta los supuestos del modelo bajo el que se construye el instrumento: a) <i>Teoría Clásica</i>: La selección de ítems de un test según esta teoría se realiza teniendo en cuenta los índices de homogeneidad, validez, dificultad y discriminación, y el coeficiente de fiabilidad del test, todo ello en relación con el número de ítems del mismo. b) <i>Teoría de Respuesta al Ítem</i>: Si el test se construye según los modelos de esta teoría, los criterios de selección de ítems serán la dificultad, discriminación, bondad de ajuste y grado deseado de la función de información del test en relación con el número de ítems.
Aplicaciones experimentales	Esta fase incluye una serie de aplicaciones experimentales a muestras más amplias que en la fase anterior, con objeto de analizar la fiabilidad y validez del test, y realizar

	un análisis más riguroso de los ítems. Determinación de las instrucciones, el tiempo de aplicación y otras características de interés del instrumento.
Aplicación definitiva y tipificación del test	El formato definitivo del test debe aplicarse a muestras grandes para ofrecer baremos adecuados y diversos. El proceso de tipificación del test consiste en una serie de transformaciones de las puntuaciones directas obtenidas de la aplicación del test, de forma que puedan ser interpretadas de manera correcta. Entre los tipos de escalas normativas se encuentran: escalas centiles, escalas de cocientes intelectuales, puntuaciones típicas, y puntuaciones típicas transformadas (escalas T, eneatis...).
Normas de aplicación	Todos los tests van acompañados de un manual que debe incluir un resumen de las distintas fases de la elaboración del test, la especificación de los baremos y las normas de aplicación. Se han de normalizar los aspectos de formato, administración, material, instrucciones y la situación de aplicación de la prueba.

Recientemente, Calero y Padilla (2004), han sintetizado del siguiente modo el proceso de elaboración de un test (Cuadro 15).

Cuadro 15. Proceso para la elaboración de un test II (Calero y Padilla, 2004)

PROCESO DE ELABORACIÓN DE UN TEST	
Pasos del proceso	Objetivos principales
Identificación del objetivo	Declaración explícita, precisa y completa del uso previsto de las mediciones: objetivo de uso y contexto de utilización.
Definición del constructo	Establecimiento de las relaciones del constructo con otros constructos o variables relevantes. Identificación de los comportamientos indicadores del constructo.
Diseño del test	Descripción detallada de la estructura del test y de los ítems.
Redacción de los ítems	Elaboración de una población de ítems.
Análisis de los ítems	Seleccionar ítems de calidad para la versión final del test. Criterios: dificultad y discriminación.
Análisis de la fiabilidad y validez	Obtener evidencias sobre la precisión y validez de las mediciones.
Normas de interpretación	Elaboración de baremos o criterios para interpretar las mediciones.

6.2.3. INTERPRETACIÓN DE LAS PUNTUACIONES UTILIZADAS

Aunque el estudio de las puntuaciones, sus características y propiedades, es propio de la Psicometría, dado que el evaluador utiliza unidades de medida, y éstas suelen representarse en puntuaciones, comentaré de forma general algunas cuestiones básicas a tener en cuenta como el valor de las puntuaciones directas, criterio y norma (ver Cuadro 16).

Cuadro 16. Puntuaciones directas, criterio y norma

PUNTUACIONES DIRECTAS, CRITERIO Y NORMA	
Puntuación directa	El evaluador atribuye números a las situaciones a evaluar y la puntuación directa es el resultado directo obtenido al evaluar el test, el número de puntos obtenidos por el sujeto. Esta puntuación no aporta en sí misma información y sólo adquiere significado si se la compara con algo, es decir, dentro de un marco de referencia.
Puntuación referida a criterio	Una puntuación se haya referida a un criterio cuando el marco de referencia en el que hay que interpretarla es un criterio de logros u objetivos a cumplir. Por ejemplo, si consideramos que los sujetos deben contestar correctamente al menos a 10 de una serie de cuestiones que les proponemos, éste sería el criterio con el que comparar los aciertos reales obtenidos por cada sujeto. Un ejemplo de una prueba referida a criterio sería un examen.
Puntuación referida a norma	Una puntuación se haya referida a una norma cuando el marco de referencia con que compararla para que obtenga significado es un grupo, es decir, se obtiene la posición del sujeto evaluado con respecto a un grupo de referencia. Hay distintos tipos, pero las más utilizadas son las puntuaciones percentiles y las puntuaciones típicas. <i>Puntuación percentil:</i> Indica el porcentaje del grupo que sobrepasa el sujeto en cuestión o que obtiene puntuaciones inferiores a este. Un percentil de 70 indica que el sujeto obtiene puntuaciones superiores al 70% de su grupo de referencia. <i>Puntuaciones típicas:</i> Indican cuanto se separa el sujeto de la media del grupo, en función de lo que se separan los demás. Es decir, la puntuación directa se interpreta en relación a la media del grupo, pero tomando como unidad de medida la desviación típica de ese grupo. No es lo mismo separarse 4 puntos de la media del grupo cuando la mayoría de los sujetos sólo se separa 2, que cuando la mayoría se separa 6.

Como acabo de comentar, los datos obtenidos en un proceso de evaluación pueden tener, al menos, un doble marco de referencia interpretativa: a) en función de las normas establecidas, y b) según unos criterios prefijados. La principal distinción entre ambas radica en la referencia comparativa tomada como elemento de contraste (Forns, 1993).

En la evaluación referida a normas se comparan los resultados del sujeto con las ejecuciones “normales” para un determinado grupo. Este es el caso de la mayoría de pruebas construidas con la metodología clásica psicométrica, donde las puntuaciones del individuo se expresan en términos cuantitativos (centiles, decatipos...). Forns (1993) distingue dos tipos de normas: la de edad o de desarrollo y la de potencia o capacidad, dentro de las cuales hay que

considerar las de ejecución máxima y las de ejecución típica. Habitualmente el análisis del desarrollo está basado en la adopción de normas de edad, y el análisis de capacidades, aptitudes o rendimientos suele basarse en normas de grupo.

Sin embargo, la evaluación criterial tiene por finalidad la descripción minuciosa y exacta de un hecho, y la comparación de los resultados del sujeto con tal esquema. Es decir, se trata de analizar aquello que hace un sujeto, su repertorio conductual, el lugar que un individuo ocupa a lo largo de las dimensiones de un hecho analizado. Los tests referidos a criterio, por lo tanto, requieren de una muy exacta definición del contenido de análisis y de una detallada especificación de la conducta.

En la misma dirección que otros autores (Forns, 1993; Knutson y Shinn, 1991) considero que el papel de los tests referidos a criterios en la evaluación es de extraordinaria importancia por dos razones. Proporcionan una importante pauta para la planificación de intervenciones terapéuticas o educativas y usadas paralelamente a pruebas normativas permiten un análisis del sujeto desde una doble perspectiva intra e interindividual. Un interesante análisis sobre el procedimiento a seguir para la construcción de pruebas criterioles ha sido realizado por Rivas, Jornet y Suárez (1995) y en el Siglo XXI muchos investigadores están trabajando en este campo. No obstante, cabe resaltar que toda evaluación es siempre normativa.

6.2.4. PROPIEDADES DE LOS TESTS Y DE LAS TÉCNICAS ESTÁNDAR DE MEDIDA

El progreso de la ciencia depende de la calidad de sus medidas de los fenómenos relevantes. Sin duda, la efectividad y valía de una evaluación descansa en buena medida en los datos que se han recogido sobre una determinada conducta; evidentemente estos deben contar con plenas garantías de objetividad y ajuste. Este asunto ha preocupado a evaluadores y tecnólogos psicometristas, que elaboraron muy tempranamente unos coeficientes de contraste a los que se han sometido todos los tests elaborados: la fiabilidad y la validez.

Estos conceptos fueron formulados primitivamente desde la teoría clásica de los tests, para posteriormente ser reformulados, ampliados y adaptados a las nuevas alternativas evaluadoras propuestas desde el enfoque conductual. El valor de la evaluación de una conducta determinada descansa en los datos que se hayan recogido y en los instrumentos empleados para recoger dichos datos, que deben ser procedimientos fiables y válidos.

6.2.4.1. LA FIABILIDAD

El término fiabilidad, según el Diccionario de la Real Academia, hace referencia a la probabilidad del buen funcionamiento de una cosa. Cuando hablamos de fiabilidad de un instrumento de medida, en el caso que nos ocupa, de un test, hacemos referencia a la precisión o grado de exactitud de la medida (Alonso Tapia, 2004a; Anastasi, 1988; Fuentes, Mateu, Miquel y García-Merita, 1989; Urbina, 2007). La fiabilidad implica consistencia y precisión (Urbina, 2007). Para Fernández-Ballesteros y Maciá (1992) la fiabilidad de una técnica de medida se define como la precisión con la que mide. Dicha precisión puede ser entendida de muchas formas pero la más común se refiere a qué medidas repetidas del mismo fenómeno producen informaciones equivalentes. Existen varias formas de abordar la fiabilidad. Fernández-Ballesteros y Calero (2004d) enfatizan la existencia de tres tipos de evidencias de fiabilidad:

1. Relacionadas con la estabilidad de la medida, es decir, con la posibilidad de que medidas hechas en momentos distintos sean iguales (test-retest).
2. Relacionadas con la objetividad del registro: que no haya diferencias entre los datos obtenidos por diferentes evaluadores (interevaluadores).
3. Relacionadas con la homogeneidad de la prueba en aquellos tests que se componen de un agregado de elementos: que todos los ítems o reactivos que componen una técnica sean equivalentes (consistencia).

6.2.4.1.1. Fiabilidad Test-Retest o Estabilidad Temporal del Test

Indica hasta qué punto se puede esperar que un individuo alcance la misma puntuación en ocasiones distintas (Fernández-Ballesteros y Maciá, 1992; Fernandez-Ballesteros y Calero, 2004). Para el cálculo de este tipo de fiabilidad es necesario repetir el test en una segunda ocasión. El coeficiente de fiabilidad en este caso es la correlación entre las puntuaciones obtenidas por los mismos sujetos en las dos aplicaciones del test, entre las que debe mediar un intervalo de tiempo no superior a 6 meses. Este tipo de coeficiente se centra en la estabilidad temporal de las puntuaciones obtenidas en el test (Anastasi, 1988; Anastasi y Urbina, 1998).

En general para considerar aceptable un coeficiente de fiabilidad se debe tener en cuenta: (a) Que su fiabilidad sea similar o mejor a la fiabilidad de tests que estén midiendo lo mismo; y (b) Que el valor del coeficiente de fiabilidad sea lo más cercano posible a 1, considerándose los valores alrededor de 0.75 como bastante buenos (Fuentes et al., 1989).

6.2.4.1.2. Consistencia Interna

Indica en qué medida los elementos que componen una prueba son homogéneos o consistentes entre sí (Fernández-Ballesteros y Maciá, 1992; Fernandez-Ballesteros y Calero, 2004). Existen varias fórmulas de estudiar la consistencia interna a partir de las covarianzas entre todos los ítems. Entre los más utilizados cabe destacar: 1) El coeficiente alpha de Cronbach; 2) La división de las dos mitades que consiste en la división del test en dos mitades lo más exactamente comparables y aplicación de la fórmula de Sperman-Brown; 3) Fiabilidad de Kuder-Richardson (KR20), cuyo cálculo se basa en la consistencia interna de las respuestas de los sujetos a todos los elementos del test. Cuanto más homogénea sea el área examinada, más alta será la consistencia entre los elementos del test.

6.2.4.1.3. Fiabilidad del Evaluador o Interjueces

Corresponde al grado en el cual lo obtenido por un observador es semejante a lo obtenido por otro/s (Fernández-Ballesteros y Maciá, 1992; Fernandez-Ballesteros y Calero, 2004). Ciertos tests, especialmente los de creatividad y los proyectivos, dejan un gran margen al juicio del examinador, o persona que evalúa el test. En estos tests la fiabilidad del evaluador puede hallarse mediante una muestra de los protocolos del test puntuados independientemente por dos evaluadores. Las dos puntuaciones obtenidas de esta forma por cada sujeto se correlacionan, y el coeficiente de correlación resultante es una medida de la fiabilidad del evaluador (Fuentes et al., 1989). La observación conductual utiliza este tipo de fiabilidad, denominándola fiabilidad del observador o fiabilidad de los jueces.

Para medir estos tres aspectos de la fiabilidad se han utilizado gran cantidad de estadísticos (correlación de Pearson, fórmulas de Kuder-Richardson, la corrección de Spearman-Brown...), cuyos resultados suelen aparecer en todos los manuales de las técnicas estándar de evaluación. El evaluador debe de analizar estos datos con la finalidad de elegir el test fiable y adecuado al caso que le ocupa.

El concepto de fiabilidad se asienta en el error de medida. Se supone que para cada dimensión existe una puntuación verdadera, la que reflejaría con exactitud lo que se pretende medir, y otra observada, que es la que nos ofrece el test. Entre ellas media el error aleatorio o de medida (Anastasi, 1988; Anastasi y Urbina, 1998; Fernández-Ballesteros y Maciá, 1992; Hayes, Nelson y Jarret, 1986; Kline, 1985; Silva, 1985). Este error procede del hecho de que los ítems que componen un test no representan todo el universo de ítems con el que se mediría un rasgo, sino que sólo nos brindan aproximaciones (que serán tanto más precisas cuanto mayor sea la fiabilidad del test). En palabras de Yela (1979), si se desconoce la fiabilidad de un test, sabemos que su medida no es precisa, pero no sabemos la magnitud de su imprecisión. Si conocemos la fiabilidad de un test, podemos averiguar, no sólo que es impreciso, sino en qué medida lo es y hasta qué punto podemos fiarnos de sus resultados.

En opinión de algunos autores (Fernández-Ballesteros, 1983; Silva, 1985) mientras que ciertos tipos de propiedad como la objetividad, son siempre necesarios, otros, como la consistencia interna o la estabilidad temporal del test pueden no serlo. No habría que buscar consistencia interna en los distintos elementos de una entrevista, porque no es probable que éstos covaríen entre sí. Un test de muy baja estabilidad puede estar indicado como instrumento valorador del cambio, en el caso de que su falta de estabilidad test-retest sea debida a su sensibilidad a los cambios ambientales. Por el contrario, deberemos seleccionar una medida que represente una alta fiabilidad test-retest si pretendemos examinar repertorios de conducta que suponemos estables (por ejemplo, en un test de inteligencia) (Fuentes et al., 1989).

Buena parte de la noción de fiabilidad, concretamente la que se refiere a la consistencia transtemporal (test-retest), se basa en el supuesto de la estabilidad de la conducta. Si la conducta presenta inestabilidad temporal, la faceta tiempo será irrelevante para la Evaluación Conductual (Barrios y Hartman, 1986).

Dentro de la tradición psicométrica se han dado discrepancias ante este supuesto de estabilidad conductual. Tallent (1965) llamó la atención sobre algunos aspectos del comportamiento que ofrecían menor estabilidad que otros. Lógicamente, los tests destinados a medir tales aspectos deberían presentar una fiabilidad test-retest más baja. El concepto de fiabilidad, en sus inicios muy impregnado de la teoría de los rasgos, fue rechazado inicialmente por los evaluadores conductuales por suponerlo emparejado con el concepto de estabilidad conductual. Este hecho ha sido criticado por algunos autores (Barrios y Hartman, 1986; Silva, 1985) que piensan que la cuestión de la estabilidad debe enfocarse bajo su prisma empírico y no teórico-inferencial.

Dado que las conductas varían en cuanto a su grado de estabilidad, la fiabilidad test-retest de un instrumento de evaluación debe ser evaluada sobre las bases de lo que se conoce acerca de la conducta concreta que se supone mide (Barrios y Hartman, 1986). Esta afirmación incluye la noción de que una escasa fiabilidad test-retest no amenaza necesariamente la bondad del instrumento, ya que puede estar explorando una conducta poco estable.

En opinión de Silva (1989) los desarrollos del pensamiento psicométrico tienden a distinguir entre fiabilidad, como propiedad de la medida, y consistencia y estabilidad como

atributos que el comportamiento puede poseer en mayor o menor grado (...), la puntuación más fiable no será la más consistente y estable, sino la que refleje la consistencia y la estabilidad u otra vertiente de generalizabilidad, que se desprenda de los atributos de aquella que se pretende medir.

Como han subrayado Fernández-Ballesteros y Calero (2004), la Teoría de Respuesta al Ítem (TRI), entiende la fiabilidad como la precisión o error cometido a la hora de estimar la puntuación de cada persona. El análisis del error de medida es muy diferente del realizado por la Teoría Clásica de los Test (TCT). La diferencia entre ambas aproximaciones es que para la TCT el concepto de fiabilidad opera globalmente para el conjunto de los valores posibles de la habilidad, y por lo tanto se define de manera incondicionada; mientras que para la TRI la función (de información) característica del ítem introduce el concepto de precisión (con la que se hace la estimación para un sujeto) que es independiente de cualquier criterio normativo, opera localmente, y se condiciona a niveles particulares de habilidad. Con la finalidad de clarificar los procedimientos de obtención de evidencias de fiabilidad y las teorías base desde las que se pueden llevar a cabo, estas autoras han elaborado un cuadro que se presenta a continuación.

**Cuadro 17. Procedimientos de obtención de evidencias de fiabilidad
(Fernández-Ballesteros y Calero, 2004)**

TIPO DE EVIDENCIA	PROCEDIMIENTO EMPÍRICO	TEORÍA BASE
Estabilidad de la medida	Formas paralelas Test-retest Estabilidad de la línea base	TCT Conductual
Objetividad del registro	Formas paralelas Test-retest Correlación entre observadores	TCT Conductual
Homogeneidad de la prueba	Dos mitades Covarianza de los ítems Análisis de varianza	TCT
Precisión	Función de información	TRI
Exactitud	Correlación con criterios objetivos	Conductual

En el contexto del debate sobre la fiabilidad, recientemente Sánchez-Meca y López-Pina (2008), subrayan que frases del tipo “la fiabilidad del test es 0.80” son incorrectas. Desde su punto de vista es más apropiado decir “la fiabilidad de las puntuaciones del test en una determinada aplicación del mismo es 0,80”. El enfoque meta-analítico de generalización de la fiabilidad pretende demostrar que la fiabilidad es una propiedad empírica que varía de una aplicación a otra del test. Este nuevo enfoque meta-analítico está contribuyendo a concienciar a los investigadores sobre la importancia de aportar estimaciones de la fiabilidad con los propios datos y evitar inducciones de la fiabilidad. En su trabajo identifican las fases en las que se lleva a cabo un estudio de generalización de la fiabilidad: 1) formulación del problema, 2) búsqueda de los estudios, 3) codificación de los estudios, 4) análisis estadístico e interpretación y 5) publicación. Además, se presenta una visión actualizada de los problemas estadísticos de este enfoque: 1) transformar versus no transformar los coeficientes, 2) ponderar versus no ponderar los coeficientes, 3) cómo tratar la dependencia estadística entre los coeficientes, y 4) cuál es el modelo estadístico más apropiado (efectos fijos, efectos aleatorios, efectos mixtos).

6.2.4.2. LA VALIDEZ

Mientras que el proceso de fiabilidad ha permanecido invariante desde los estándares de 1954 (APA), el concepto de validez con el paso del tiempo ha sufrido grandes cambios (Fernández-Ballesteros y Calero, 2004).

Desde la definición clásica la validez de un instrumento de medida hace referencia al grado en que este mide en realidad lo que pretende medir (Fuentes et al., 1989; Urbina 2007). Se dice que un instrumento es válido cuando mide lo que pretende medir y no otra cosa (Vernon, 1964) o como dice Kaplan (1964) cuando se muestra hábil para cumplir los objetivos específicos señalados por el evaluador. La APA (1974) define la validez como el grado en que pueden ser sostenidas las inferencias a partir de los datos de la evaluación, ya que la validez, o adecuación consigo misma, no es una entidad mensurable, también representa una inferencia. La validez del test se refiere a lo que éste mide y a cómo lo mide (Anastasi, 1968/1978), tiene que ver con lo que miden y con qué tan bien lo hacen (Anastasi y Urbina, 1998), y también se refiere a la comprobación empírica de que el instrumento de que se trate mida la variable que pretenda medir. Como se sabe, todo instrumento válido ha de ser fiable, pero no toda medida fiable es válida (Fernández-Ballesteros y Maciá, 1992).

El problema de la validez de un instrumento de medida importa a todas las ciencias, pero se plantea con especial necesidad y urgencia en las humanas. Parece claro que el metro sirve para medir la longitud, y el termómetro para medir la temperatura. Pero ¿Qué mide en el fondo un test llamado de inteligencia? ¿Mide la inteligencia, la cultura, el grado de conocimientos o la manera de pensar de un cierto grupo de personas? (Yela, 1977).

Lo mismo que hemos dicho al referirnos a la fiabilidad de las técnicas de evaluación, es aplicable a la validez. Es decir, debemos seleccionar nuestros instrumentos generales en base a los datos que, sobre su validez, aparecen en los manuales correspondientes. Esto proporciona garantías científicas a nuestro trabajo (Fernández-Ballesteros y Maciá, 1992; Fuentes et al., 1989).

Lo primero que debe hacerse para comprobar la validez de un test es definir con la mayor exactitud y rigor posible el fin al que se destina, es decir, el criterio de validez (fin que se pretende conseguir con el test). Se denomina coeficiente de validez a la correlación entre las puntuaciones de los sujetos en el test y sus puntuaciones en el criterio de validez. Es poco corriente que un coeficiente de validez suba más allá de 0,60, no obstante, siempre hay que intentar conseguir el coeficiente más alto posible.

Aiken (1996) subrayó que una prueba puede caracterizarse por muchos tipos de validez, en función de los propósitos específicos con los que se diseñó, de la población a la que se dirige y del método para determinar la validez. Los tests pretenden servir a muchos fines y de maneras muy distintas, por ello se han elaborado diferentes conceptos en torno a la validez. Estos conceptos han sido descritos como tipos de validez, como si fueran independientes. En este sentido desde la teoría clásica de los tests (TCT) se habla de tres tipos de validez: Validez de Contenido, Validez Criterial, y Validez de Constructo. Todos ellos se refieren a lo mismo, todos expresan la medida en que el test logra lo que se propone, pero cada uno subraya un matiz distinto.

6.2.4.2.1. Validez de Contenido (Muestral)

El concepto de validez de contenido se refiere al grado en que los ítems de un instrumento comprenden una muestra representativa de la conducta que se pretende medir (Anastasi, 1988). Supongamos que queremos medir los conocimientos sobre Historia Contemporánea que tienen un grupo de alumnos, para lo que preparamos una prueba con 25 preguntas, de las cuales 20 hacen referencia a la Guerra Civil Española. Como es evidente nuestra prueba no mide los conocimientos sobre historia contemporánea (que comprende desde la Revolución Francesa hasta nuestros días), sino más bien sus conocimientos sobre la Guerra Civil Española, por lo que esta prueba no tiene validez de contenido o muestral en el área de los conocimientos que nos interesaba (Fuentes et al., 1989).

Según el criterio de la APA (1974, 1976, 1981), se requiere evidencia sobre la validez de contenido cuando quien utiliza el test desea estimar el rendimiento de un individuo en el universo de situaciones que el test intenta representar (Silva, 1989). A juicio de Cone (1981) la validez de contenido sólo ha sido contemplada desde la perspectiva tradicional por los usuarios de los tests de rendimiento (miden el grado de dominio del individuo en una habilidad específica o en un curso de estudio), sin embargo, es la más consensuada entre los evaluadores conductuales (Fernández-Ballesteros y Maciá, 1992; Silva, 1989). Esto es explicable por la estrecha relación con el enfoque referido a criterios, propio de la Evaluación Conductual.

6.2.4.2.2. Validez Criterial (Empírica)

El concepto de validez de criterio expresa el grado en que las puntuaciones de una variable (predictora) sirven para inferir el rendimiento en otra variable (criterio). "Criterio" es, pues, la variable que debe ser predicha y "predictora" aquella a través de la cual se predice. Este criterio es externo al test y puede ser, por ejemplo, otro test, una estimación de jueces, o el promedio de una serie de puntuaciones obtenidas por el individuo en diversas pruebas o situaciones (Pelechano, 1988d). El grado de relación se determina empíricamente mediante métodos estadísticos de correlación, y se denomina coeficiente de validez.

En cada caso, el problema estriba en elegir un criterio, lo más adecuado posible, y medirlo con precisión. Fijado este criterio y conseguida su medida, se averigua la correlación entre el test y el criterio. Este coeficiente de validez expresa el grado en que el test es empíricamente válido para apreciar el criterio (Yela, 1977). Si un test de inteligencia infantil, por ejemplo, tiene alta correlación con el éxito escolar, es un test empíricamente válido para distinguir a los que tienen éxito escolar de los que no lo tienen.

Para Anastasi (1988) la validez de criterio indica la eficiencia de un test para predecir la conducta de un individuo en situaciones específicas. Para Fernández-Ballesteros y Maciá (1992) la validez criterial establece la medida en la cual el instrumento está asociado a un criterio independiente. Este tipo de validez empírica puede ser concurrente o propiamente predictiva, diferenciándose ambas por las relaciones temporales test-criterio.

La validez de criterio indica la utilidad de las medidas proporcionadas por los test. El criterio puede ser cualquier conducta que se desee predecir a partir de las puntuaciones observadas en el test. Se trata, por tanto, de validez en relación con un criterio. Se distinguen dos formas según el tiempo que existe entre la aplicación del test y la medición del criterio: validez concurrente y validez predictiva.

Validez Concurrente: Se denomina así a la validez criterial que es estimada obteniendo puntuaciones al mismo tiempo (o en tiempos muy próximos) en el test y en el criterio. El coeficiente de validez es el coeficiente de correlación entre las puntuaciones obtenidas en el test y el criterio.

Validez Predictiva: Indica el grado en que las puntuaciones de un test predicen conductas futuras. El coeficiente de correlación obtenido entre el test y el criterio, tras un periodo razonable de tiempo, equivale al coeficiente de validez predictiva.

La validez concurrente hace referencia a la amplitud de la correlación entre medidas parecidas de un mismo criterio. La validez predictiva de un instrumento de evaluación viene dada por su potencia predictiva de algún criterio (Fernández-Ballesteros y Maciá, 1992; Kline, 1985; Silva, 1989). Por ejemplo, un test que tenga validez predictiva nos permite contestarnos a preguntas del tipo: ¿Es posible que Paco se vuelva neurótico? ¿Es posible que Pedro sea un buen conductor? Mientras que un test que tenga validez concurrente nos da contestación a las preguntas de tipo: ¿Está Paco neurótico? ¿Es Pedro un buen conductor? (Fuentes et al., 1989).

Desde la perspectiva conductual la validez criterial ha sido considerada por muchos como irrelevante, al compartir la creencia de que el instrumento de evaluación constituye de por sí un buen criterio, sobre todo cuando se trata de la observación directa en medios naturales. Sin embargo, la observación directa no siempre es posible. Para Silva (1989) es necesario recurrir a la validez criterial en las siguientes situaciones: a) en los casos en que sea necesario realizar una evaluación indirecta; b) cuando se utilizan unidades de análisis de conducta muy complejas; y c) en la predicción sobre los resultados de la intervención.

El fracaso de un número importante de estudios predictivos ha ido generando una mayor preocupación por la validez conceptual de los procedimientos empleados, lo que lleva a ver la validez criterial a través de la óptica de la validez de constructo, estableciéndose entre ambas una relación jerárquica. Silva (1989) afirma que dentro del marco del concepto unitario de la validez... se trata de interpretar los problemas de predicción en función de los procesos psicológicos que están en la base y que permiten una explicación científica de la misma. Esta concepción más actual significa englobar las cuestiones de validez criterial dentro del concepto más amplio, y jerárquicamente superior, de validez de constructo.

En una dirección similar, Messick (1994) definiendo la validez como un juicio evaluativo de acuerdo integrado para el cual la racional evidencia empírica y teórica apoya la adecuabilidad de las interpretaciones y acciones basadas en las puntuaciones de los tests u otros modelos de evaluación, considera que la tradicional concepción de validez dividida en tres tipos separados (contenido, criterio y constructo) es una concepción fragmentada e incompleta, defendiendo un concepto unificado de validez que interrelaciona estos temas como aspectos más o menos en la teoría de la validez de constructo.

6.2.4.2.3. Validez de Constructo (Teórica)

La validez de constructo, también llamada estructural o conceptual, establece el grado en el cual una técnica mide un determinado rasgo o constructo hipotético (Fernández-Ballesteros y Maciá, 1992). Fue desarrollada por Cronbach y Meehl (1955) y tuvo gran relieve dentro de la teoría del rasgo. Se utiliza en los casos en que el examinador no dispone de criterios bien

delimitados y tiene que servirse de medidas indirectas o constructos, relacionados con una determinada teoría. El constructo que subyace al test será validado en la medida en que se mantengan relaciones entre las puntuaciones del test y las características del constructo predichas por la teoría que lo sustenta (Anastasi, 1988; Fernández-Ballesteros y Maciá, 1992; Kline, 1985; Silva, 1989).

Campbell y Fiske (1959) ampliaron los conceptos y métodos pertinentes con la validez de constructo mediante la introducción de dos términos complementarios: la "validez convergente" y la "validez divergente o discriminante". El método empleado para estudiar estas modalidades de validez es denominado "matriz multirasgo-multimétodo".

Validez Convergente: Expresa el grado en que una medida está sustancialmente relacionada con otras medidas del mismo constructo y con otras variables con las que, por razones teóricas, debería relacionarse.

Validez Divergente o Discriminante: Expresa el grado en que una medición presenta una baja correlación con otras mediciones que, utilizando el mismo método, miden otros constructos.

Mientras que algunos autores critican el concepto validez de constructo, otros tratan de revisarlo y adaptarlo a la perspectiva conductual. Silva (1989) hace una detallada revisión del concepto de validez de constructo que resume en 10 puntos:

1. Constructo es sinónimo de concepto científico y validez de constructo, de validez conceptual o grado de adecuación de las inferencias conceptuales o teóricas que se hacen a partir de los datos de evaluación.
2. Los constructos no deben ser considerados como algo estático.
3. Los constructos aparecen relacionados, generalmente de forma probabilística, con indicadores observables y con otros constructos, constituyendo redes nomológicas.
4. Un modo de explorar la entidad del constructo y su red nomológica es a través de la validez convergente y divergente.
5. Toda validación de constructo se refiere tanto al concepto como al método implicado.
6. El status de los constructos en el sistema científico implica: a) tanto los constructos como su validación están ligados a la evidencia empírica; b) un constructo no se reduce a sus referentes empíricos; c) el constructo posee un status fundamentalmente epistemológico; d) los constructos se proponen no sólo con fines especulativos, sino también de predicción; y e) el valor de un constructo se juzga por su utilidad.
7. La validez de constructo engloba a la de contenido y a la criterial. Se confunde, por tanto, con el concepto de validez.
8. No existe límite en cuanto a las estrategias, procedimientos o instrumentos potencialmente útiles en una validación de constructo.
9. Al proceso de validación de constructo no puede ponerse término, en la medida en que nunca puede considerarse completo.
10. La validación de constructo consiste esencialmente en la aplicación del proceso de formulación y contrastación de hipótesis científicas al campo de la Evaluación Psicológica.

El concepto de validez de constructo es, por tanto, más inclusivo que el de validez de criterio. La diferencia entre ambos conceptos estaría en la amplitud conferida a la variable criterio que se maneja en uno u otro caso. Sin embargo, Silva (1989) entre otros autores, estima que el concepto de validez es esencialmente unitario y que los tipos de validez deben considerarse, más bien, como componentes de validez: Las categorías de criterio, contenido y constructo, pueden ser tan convenientes como otras, no implican distinciones tajantes, no deben utilizarse como tipos de validez ni se corresponden puntualmente con tipos separados de inferencias o de utilización de los instrumentos (Silva, 1989).

Después de realizar una revisión del concepto de validez, Silva (1989) hace las siguientes precisiones:

- La validez está relacionada con las inferencias que se hagan a partir de las puntuaciones obtenidas mediante un instrumento en determinadas circunstancias.
- En sentido estricto, no se valida al instrumento, sino a las interpretaciones que se hacen a partir de sus puntuaciones.
- La validez es algo que se infiere a partir de un conjunto de informaciones y no es algo que se reduce a un coeficiente.
- No debe hablarse de tipos de validez. El concepto de validez es esencialmente unitario.
- No existe límite en cuanto al procedimiento o de datos para estimar la validez. Cualquier información puede ser relevante en un proceso de validación, el cual es identificado con el proceso científico de construcción y contrastación de hipótesis.

Los problemas que plantea la visión tradicional o clásica de la validez giran en torno a ciertos supuestos implícitos: 1) la validez es una propiedad de los tests, más que de las interpretaciones de las puntuaciones de los tests; 2) para que fuesen válidas, las puntuaciones de los tests deberían medir directamente algún constructo predeterminado; y 3) La validez de una puntuación es, al menos en parte, función de la manera en que el autor del test entiende el constructo que intenta medir (Urbina, 2007).

6.2.4.2.4. Últimos desarrollos del concepto de validez: Estándares del APA (1999)

Los tests psicológicos existen para ayudarnos a hacer inferencias sobre la gente y la conducta. La validez se basa en la seguridad con que pueda apoyarse cualquier inferencia que se haga a partir de los resultados de los tests (Urbina, 2007). La primacía de las consideraciones sobre la validez se reconoce en las actuales normas sobre el uso de los tests, que define la validez como “el grado en que las pruebas acumuladas apoyan que la interpretación propuesta a partir de las puntuaciones del test es adecuada para la finalidad pretendida” (AERA, APA, NCME, 1999).

Los estándares del APA (1999) presentan una teoría unificada de la validez en la que el constructo es el eje central de la validación y en la que incorporan el énfasis en las consecuencias del uso de los test, el papel de la teoría de la generalizabilidad y el uso de los test para la toma de decisiones políticas, y definitivamente aportan la siguiente definición: “Validez es un concepto unitario. Es el grado en que todas las evidencias acumuladas apoyan la proyectada interpretación de las puntuaciones del test implicadas en el propósito propuesto... Validez se refiere al grado en que evidencia y teoría soportan las interpretaciones de las puntuaciones del test implicadas en los usos propuestos del mismo... El proceso de validación

implica acumulación de evidencias que proporcionan una base científica sólida para las interpretaciones que se proponen. Esto es, deben ser evaluadas las interpretaciones de las puntuaciones del test requeridas para usos propuestos, no el test mismo”.

Así, el planteamiento general de los estándares de 1999, vigentes en la actualidad, no sólo reafirma la posición central de la validez, sino que entiende que los principios de la validez se basan en las inferencias sobre el significado de comportamientos o de atributos consistentes. Las principales fuentes de evidencias son (Fernández-Ballesteros y Calero, 2004):

1. Evidencias basadas en el contenido del test: incluye evidencias de relevancia, representatividad y calidad técnica del instrumento. Estas proceden del análisis de las relaciones entre el contenido del test y el constructo que se pretende medir. Por contenido se entiende los temas, palabras, formatos de ítems y/o tareas que forman el test tanto como las instrucciones y los procedimientos de administración y puntuación.
2. Evidencias basadas en los procesos de respuesta (validez sustantiva): en éstas se trata de recurrir al análisis teórico y empírico de los procesos de respuesta de las personas que cumplimentan el test para valorar el ajuste entre el constructo y la respuesta real de los examinados.
3. Evidencias basadas en la estructura interna: tratan de analizar las relaciones de los ítems de test entre sí y de ellos con el constructo que sirve para la interpretación de las puntuaciones obtenidas.
4. Evidencias basadas en la relación con otras variables (validez externa): se incluyen aquí los análisis de relaciones entre las puntuaciones del test y otras variables externas utilizadas como criterio, sean ejecuciones o puntuaciones de otros tests que midan el mismo constructo o constructos diferentes.
5. Evidencia de generalizabilidad serían aquellas que examinan la extensión en que las puntuaciones y su interpretación son generalizables a través de grupos, contextos y tareas incluyendo relaciones test-criterio.
6. Evidencias sobre las consecuencias del uso del test: se trata de atender a las consecuencias previstas y no previstas del uso del test, potenciales y reales, positivas y negativas.

En definitiva, parece que con los estándares de 1999 se ha llegado a la reunificación de un concepto de validez aplicable a las diversas técnicas que pueden utilizarse en evaluación. Con esta definición de validez:

1. Se recogen las teorías y métodos que durante estos años se han desarrollado: la teoría de la generalizabilidad, métodos factores en las matrices multirrasgo x multimétodo, teoría de respuesta al ítem...
2. Se incorpora, además, el interés por lo social y por la responsabilidad del psicólogo en relación a la toma de decisión que él u otros, según sus datos (políticos, por ejemplo), hacen.
3. Se integran conceptos que, aunque respondían al nombre de validez, se separaban su consideración de requisito psicométrico, pues se asociaban más a efectos del proceso de evaluación que a las técnicas en sí. Éstos eran los conceptos de validez aparente, validez social y/o utilidad.
4. Implica la aceptación del concepto de validez desde esquemas conductuales de construcción de técnicas de evaluación.

Con la finalidad de clarificar los procedimientos de obtención de evidencias de validez y las teorías base desde las que se pueden llevar a cabo, Fernández-Ballesteros y Calero (2004) han elaborado un cuadro que se presenta a continuación.

**Cuadro 18. Procedimientos de obtención de evidencias de validez
(Fernández-Ballesteros y Calero, 2004)**

TIPO DE EVIDENCIA	PROCEDIMIENTO EMPÍRICO	TEORÍA BASE
Evidencias basadas en el contenido	Análisis de temas, palabras, formatos de los ítems y/o códigos conductuales Instrucciones para aplicadores u observadores Métodos de muestreo Directrices de administración y puntuación	TCT Conductual
Evidencias basadas en los procesos de respuesta	Análisis de los procesos de respuesta Datos que muestran que el sujeto responde utilizando procesos adecuados Métodos de entrenamiento para eliminar sesgos en los observadores	TCT Conductual
Evidencias sobre la estructura interna	Análisis factoriales Análisis de consistencias	TCT
Evidencias de generalizabilidad	Correlación con criterios Precisión Exactitud	TCT TRI Conductual
Evidencias basadas en la relación con otras variables	Evidencias convergentes y discriminantes, modelos estructurales Red nomológica Exactitud	TCT TRI
Evidencias sobre las consecuencias	Consecuencias para el evaluado Sensibilidad, especificidad, poder predictivo y utilidad	TCT Conductual

6.2.5. LA TEORÍA DE LA GENERALIZABILIDAD. TG.

Como señalan Fernández-Ballesteros y Calero (2004), una cosa es que los instrumentos estándar de evaluación presenten las garantías científicas necesarias para basar en ellos nuestras inferencias científicas y otra que los datos recogidos en nuestra evaluación las presenten. Desde luego, no cabe duda de que cuanto mejores instrumentos utilicemos, tanto más adecuadas garantías presentarán nuestros datos. La conceptualización sobre las garantías

científicas que deben ofrecer los datos, recogidos en el proceso evaluador, procede de la teoría de la generalizabilidad (TG) formulada por Cronbach, Gleser, Nanda y Rajaratman (1972).

Inicialmente los evaluadores conductuales rechazaron los criterios psicométricos clásicos considerando que no eran aplicables a este modelo, sin embargo, progresivamente un amplio grupo de evaluadores conductuales emplean formas parecidas para contrastar la bondad de sus evaluaciones. Ante la necesidad de validar sus técnicas e instrumentos, los evaluadores conductuales han adoptado diferentes alternativas tales como: empleo de conceptos psicométricos clásicos, reformulación conceptual de estos criterios tradicionales y acercamiento de los dispositivos de contraste de los paradigmas científico-naturales.

Mientras que un sector de autores conductuales ve factible un acercamiento remodelado a los criterios psicométricos tradicionales para contrastar la bondad de sus instrumentos, otros siguen pensando que los dispositivos de contraste procedente de la psicometría clásica son algo limitados y no tienen aplicabilidad óptima a las técnicas e instrumentos de la evaluación de la conducta. Para remediar esta dificultad surge la alternativa de la Teoría de la Generalizabilidad.

Como ya se ha anticipado en los apartados previos de este capítulo, la teoría de la generalizabilidad (Cronbach et al., 1972) supone una reconceptualización más amplia de los conceptos de fiabilidad y validez que los manejados desde la psicometría clásica. Tanto los conceptos sobre fiabilidad como los de validez pueden ser entendidos a través de la teoría de la generalizabilidad, en base al grado en el cual las puntuaciones halladas mediante un instrumento de medida son generalizables a distintos ámbitos o universos de generalización. En base a ello, los autores proponen una ordenación de las distintas formas de fiabilidad y validez, y concretamente, a través de un continuo de creciente generalizabilidad.

Se sustituye el concepto clásico de “puntuación verdadera” por el de “puntuación universo”. Con este concepto se expresa el grado de inferencia que el examinador efectúa desde una muestra de datos observados a un conjunto de datos de interés procedentes de diferentes ámbitos. Los datos que nos aporta una determinada técnica de evaluación tienen interés por cuanto son muestras representativas del universo de datos que podrían ser obtenidos. Lo que interesa ante todo no es llegar a una puntuación verdadera sino fijar hasta qué punto una observación puede generalizarse a otras observaciones o universos.

En este contexto, Cone (1977) ha propuesto hasta seis universos de generalización a través de los cuales deben ser generalizados los datos de una Evaluación Conductual:

1. *Universo de generalización de las puntuaciones*: Indica la medida en que los datos obtenidos por un observador son generalizables a los obtenidos por otro (equivale al concepto de fiabilidad interjueces de la psicometría clásica).
2. *Universo de generalización de los elementos*: Indica el grado en que los ítems de un instrumento comprenden una muestra representativa del universo de conductas que se pretenden medir (equivale a validez de contenido y consistencia interna).
3. *Universo de generalización temporal*: Indica el grado en que los datos obtenidos en un momento determinado son inferibles a otros momentos de la vida del sujeto (equivale al concepto fiabilidad test-retest).
4. *Universo de generalización situacional*: Indica el grado en que los datos obtenidos en una situación son generalizables a otras (equivale al concepto de validez ecológica).

5. *Universo de generalización de los métodos*: Indica el grado en que los datos obtenidos mediante un método de evaluación son generalizables a otros que evalúan el mismo contenido (equivale al concepto de validez convergente).
6. *Universo de generalización de las dimensiones*: Indica el grado en que las estimaciones de una conducta son generalizables a otras conductas (equivale al concepto de validez de constructo).

La elección de uno u otro universo dependerá de los objetivos concretos que busquemos mediante la evaluación. Dados los objetivos fundamentales de la evaluación de rasgos (clasificar y predecir) la cuestión de la estabilidad temporal es la de mayor relevancia. Evidentemente las metas de la evaluación conductual son distintas y, por tanto, esta perspectiva se amplía desde la teoría de la generalizabilidad.

Para Cronbach et al. (1972) los universos de generalizabilidad pueden dividirse en dos grandes tipos:

1. *Universo Personas*: Hace referencia a la generalización inferencial de muestra a población. Este concepto incluye la problemática propia de todo muestreo e inferencia estadística. En cuanto a la fiabilidad, significa que los resultados para una determinada población deben basarse en muestras representativas.
2. *Universo Condiciones*: Hace referencia a un amplio repertorio de condiciones, tales como, los ítems que componen un test, las distintas formas de un test, diferentes estímulos, observadores, situaciones de medición, etc. Cada una de estas condiciones puede constituir un universo a través del cual se intenta generalizar.

Algunos autores (Fernández-Ballesteros y Maciá, 1992; Fernández-Ballesteros y Calero, 2004; Hayes, Nelson y Jarret, 1986; Márquez, 1987; Silva, 1985, 1989) han visto en esta teoría una alternativa adecuada para garantizar la bondad de los datos, ya que permite ampliar, reformular y reinterpretar los conceptos clásicos sobre validez y fiabilidad, que anteriormente tenían poca aplicabilidad en la Evaluación Conductual.

A este respecto, Silva (1985) afirmó que dos son las maneras como fiabilidad y validez están íntimamente relacionadas: en primer lugar, por pertenecer ambas a un continuo de generalizabilidad donde la fiabilidad es la expresión mínima (intratest) y la validez la expresión máxima. Pero también, y en segundo lugar no menos importante, porque las cuestiones de fiabilidad dependen de las cuestiones de validez, y en particular de la validez de constructo. El mismo autor, unos años más tarde señaló como otra de las aportaciones de la “teoría de la generalizabilidad” el hecho de que permite lanzar un puente teórico entre el concepto de fiabilidad y validez: “En efecto, si el universo “ítems” o el universo “tiempo” eran clásicamente tratados en el capítulo de fiabilidad, el universo “situaciones”, por ejemplo, lo era en el capítulo de validez (concurrente o predictiva, según los casos)... Los límites (entre fiabilidad y validez) se diluyen. Los problemas de fiabilidad y validez son (al menos en gran parte) problemas de generalizabilidad (Silva, 1989).

Un análisis de los universos de generalizabilidad puede encontrarse en el reciente trabajo de Fernández-Ballesteros y Calero (2004), en el que profundizan en los 6 universos: generalizabilidad de las puntuaciones, generalizabilidad de los elementos, generalizabilidad

temporal, generalizabilidad de las situaciones, generalizabilidad de los métodos, y generalizabilidad de las dimensiones.

6.2.6. LA TEORÍA DE LA RESPUESTA AL ÍTEM. TRI.

El campo de la medida ha avanzado de forma muy significativa durante los últimos años y ha supuesto un giro muy importante tanto en los modelos de medida como en el campo práctico de la evaluación. Un cambio fundamental ha sido la transición durante la década de los 80, desde la Teoría Clásica de los Tests (TC) a la Teoría de Respuesta al Ítem (TRI). La Teoría de Respuesta al Ítem ha reemplazado a la Teoría Clásica de medida, como marco para el desarrollo de tests, construcción de escalas, etc.

En los últimos veinte años se está produciendo una intensa reflexión acerca de las características de la medida psicológica. La teoría clásica de los tests, sustentada sobre los conceptos de muestreo, estandarización y de fiabilidad-validez de la medida, apoyada en referencias grupales, ha sufrido importantes cambios.

La crítica básica que genera la búsqueda de nuevos sistema de medida es que, tanto en la teoría clásica de los tests como en la teoría de la generalizabilidad, las puntuaciones de un test son más dependientes de la muestra que de la propia función analizada. En la psicometría clásica, la selección del ítem por los criterios de dificultad o por su poder discriminativo (parámetros de ítem) depende de su relación con las puntuaciones obtenidas por el grupo de sujetos en el resto de ítems de la prueba. El parámetro referido a la persona (habilidad) depende también de la muestra y del conjunto de ítems del test que se les ha aplicado. Las respuestas a los ítems no pueden ser interpretadas como respuestas puras de la medida pretendida, sino como un reflejo de las propiedades del conjunto de sujetos (medidas relativas) ante una tarea determinada (Forns, 1993).

La expresión “Teoría de Respuesta al Ítem” (TRI) se refiere a una amplia y creciente variedad de técnicas utilizables en el diseño y elaboración de nuevos tests y en la evaluación de los ya existentes y que son variadas por las fórmulas matemáticas que emplean, por el número de rasgos o dimensiones aptitudinales que analizan como objetivos de medida. También se utilizan diferentes métodos dependiendo de si los datos del ítem son dicotómicos (acierto-fallo, verdadero-falso...), o politómicos (múltiples opciones de respuesta). Los procedimientos que abarca la TRI son amplios y complejos y hasta hace poco las publicaciones existentes sobre estos métodos eran demasiado difíciles de comprender sin unos sólidos conocimientos en matemáticas y estadística. No obstante, en los últimos años se han publicado numerosos materiales sobre las técnicas TRI más accesibles (Urbina, 2007).

La TRI trata de responder a dos problemas fundamentales presentes en la Teoría Clásica de los Tests (TC). En primer lugar, el error en la medida. Desde la TC se asume que las puntuaciones de los sujetos en un test estarán afectadas por un error aleatorio, atribuible a diversas causas: dependientes del sujeto, del ambiente, del instrumento de medida y del propio proceso de evaluación. A lo largo de la historia se han realizado diversos intentos de clasificación del error en función de las fuentes de donde procede, siendo la teoría de la generalizabilidad de Cronbach et al. (1972) el intento más ambicioso. El segundo problema que trata de resolver la TRI se refiere a la invarianza de las mediciones y las propiedades de los instrumentos de medida.

Las nuevas propuestas de medida que pretenden resolver algunos de estos aspectos se basan en la Teoría de Respuesta al Ítem (TRI) estrechamente vinculada a la teoría de rasgo latente, desde las que se considera que es posible derivar medidas o estimaciones de rasgos latentes los cuales son identificables a partir de un conjunto de observaciones. La expansión de la Teoría de Respuesta al Ítem ha implicado derivaciones aplicadas relevantes como son: los bancos de ítems, los tests a medida del sujeto o tests adaptados y la comparación de la precisión de conjuntos de medidas en distintos momentos de una escala de rasgo latente (López Feal, 1990; Losada y López-Feal, 2003; Muñoz, 1990, 1997; Navas, 2001; Renom, 1990, 1993). Esta concepción de medida tiene los siguientes objetivos (Forns, 1993):

- La búsqueda de medidas que sean independientes de las puntuaciones estándar derivadas del grupo (medidas invariantes en relación con la muestra o medidas invariantes en relación con los examinados).
- La elaboración de nuevas pruebas que analicen la invarianza de la conducta en sí misma, de modo que un test represente con precisión un dominio gradual de conocimiento relativo a una única medida.
- La relación de los dos aspectos anteriores permite un tipo de medida en la que los parámetros de ítem y de persona son ambos invariantes, de tal modo que ni la elección de una muestra de sujetos, ni la elección de los ítems afecte a los parámetros de dificultad del ítem ni a los de la habilidad.
- La agilidad en la combinatoria de ítem de tests, que pertenezcan a un mismo dominio de conducta, dando paso a la aplicación de tests adaptados al sujeto, en función de la capacidad de las habilidades de cada individuo.

Además, en opinión de Forns (1993), otra de las ventajas de este sistema de medida es que va a permitir el diseño de perfiles individuales que especifiquen con precisión las áreas de consistencia o de debilidad de los individuos, lo cual va a sustentar las decisiones en algunas variables y va a aportar mayor conocimiento de causa cuando sea preciso el diagnóstico con finalidad clasificatoria, por ejemplo, en el dominio de la selección.

La TRI ofrece un modelo matemático logístico para describir la relación entre el nivel de habilidad del examinado y la probabilidad de que éste dé una respuesta correcta a un ítem del test (Muñoz, 1990, 1997). La probabilidad de determinada respuesta depende de un cierto número de parámetros. Dos son los parámetros básicos: (1) Parámetro persona (toma el valor del nivel de habilidad o conocimiento, según lo cual es posible describir a una persona con independencia de la muestra de ítems aplicados a tal persona); y (2) Parámetro de ítem (nivel de dificultad, por lo cual se asume que se puede describir un ítem independientemente del conjunto de sujetos que respondan a tal ítem). Así, el modelo propuesto da cuenta de cómo los sujetos de distinto nivel de habilidad (o de conocimiento) responden diferencialmente a los ítems de un test.

Dentro de la TRI existen varios modelos alternativos diferenciados en el número de los parámetros que consideran necesarios para explicar la función. Los modelos más utilizados son: (1) *El modelo de Rasch*, basado en un único parámetro, el denominado parámetro de dificultad, sitúa al ítem en una escala de dificultad, considerando que la probabilidad de éxito en la solución de un ítem depende de la habilidad de los sujetos y de la dificultad del ítem, sin que intervengan otras variables; (2) *Los modelos de dos dimensiones*, desde los que se consideran

dos parámetros, el de dificultad del ítem y el parámetro de discriminación, puesto que algunos ítems pueden ser mejores para discriminar a los sujetos de bajo y alto nivel; y (3) *El modelo de los tres parámetros*, que integra un tercer parámetro, el de la conjetura del ítem, es decir, la probabilidad de que un sujeto de baja habilidad acierte con la respuesta por casualidad.

Los más populares de los modelos de respuesta al ítem para tests de elección múltiple son el modelo de un parámetro de Rasch y el modelo de los tres parámetros. Mientras que algunos autores han sido críticos con la utilización de modelos multi-parámetros, otros como Hambleton (1994) argumentan en la dirección de continuar la investigación aplicando modelos de TRI multiparámetros.

Un concepto de gran relevancia en la TRI es el de *la curva característica de un ítem* (CCI) que es la representación gráfica de la probabilidad de dar una respuesta correcta a un ítem como una función del nivel del atributo medido por el ítem en cuestión. El atributo medido, que subyace en el nivel de realización demostrado, toma el valor de rasgo latente. Otra característica fundamental de los modelos de TRI es *el supuesto de independencia local*, es decir, el acierto de un sujeto en un ítem no viene afectado por su ejecución en los otros ítems del test. Sin embargo, como señala Forns (1993), este supuesto es difícilmente aceptable en los contextos educativos y clínicos ya que el aprendizaje humano depende de la experiencia y de la práctica.

La revisión operada por distintos autores (Aguado et al., 2000, 2005; Forns, 1993; Muñiz y Hambleton, 1992; Hontangas et al., 2000; Navas, 1994, 2001; Ozer y Reise, 1994) en relación a las aplicaciones de la TRI, coinciden en mencionar:

6.2.6.1. Los bancos de ítems

Un banco de ítems no es más que un conjunto de elementos que miden un rasgo, campo o área de conocimiento, almacenados en un ordenador junto con sus propiedades psicométricas, junto con sus parámetros (Navas, 1994). Un banco de ítems contiene los ítems parametrizados, es decir, calibrados en la misma escala, que definen operativamente la variable medida. Un banco de ítems es el conjunto de ítems que cumplen el requisito de definir con precisión la variable habilidad que se analiza. Un banco de estas características corresponde al concepto teórico, ya que cada uno de los ítems representa aspectos conceptuales de la teoría que se desea poner a prueba más que los principios clásicos de la representatividad de la conducta de una muestra. Desde este enfoque, el rasgo es el banco y será la validez de contenido del banco la que determinará el grado de generalización de los resultados (Muñiz y Hambleton, 1992; Muñiz 1997). La máxima utilidad de un banco de ítems es que es móvil y puede ser incrementado a partir de investigaciones diversas, por lo tanto, permite la posibilidad de combinar los ítems del modo que sea oportuno, creando tests más fáciles o difíciles, etc.

En la actualidad existen numerosos Bancos de Ítems (BI) de acceso internacional para varios campos profesionales y académicos. Se han construido bancos de ítems en sectores tan distintos como las Fuerzas Armadas, grandes compañías de tests y en importantes organizaciones médicas y relacionadas con la salud, pero donde se ha observado el mayor desarrollo ha sido en el campo educativo. En un mundo cada día más unificado y global los BI internacionales pueden jugar un importante papel en la extensión de los conocimientos científicos.

El programa MICROCAT para PC permite la construcción de BI, la aplicación por pantalla y la corrección y análisis de los resultados, tanto desde el punto de vista clásico (ITEMAN),

como de la TRI (RASCAL, ASCAL). Nótese que cuando se habla de invarianza de las mediciones respecto a los tests utilizados se está refiriendo a tests compuestos por ítems pertenecientes a un banco y calibrados en la misma escala, de lo contrario no hay tal. Puede decirse que sin la existencia de un banco de ítems la TRI no reporta ninguna ventaja fundamental frente a la TC (Muñiz y Hambleton, 1992).

6.2.6.2. Construcción de tests a medida del sujeto o Test adaptados computerizados

Los tests a medida fueron otro de los conceptos psicométricos potenciados por la onda expansiva de la TRI (Muñiz y Hambleton, 1992). Los denominados Tests a Medida o Tests Adaptivos han tomado fuerza ante los avances informáticos y de la TRI, derivando en los actuales Test Adaptativos Computerizados o Informatizados (TAC - TAI) (Aguado et al., 2005; Díaz Esteve, 1994; Renom, 1993, 1994). Son un corolario de la existencia de un banco de ítems, consistiendo en la selección informatizada de aquellos ítems que mejor permiten medir la habilidad de un individuo concreto; son seleccionados a partir de las respuestas del individuo a los ítems precedentes (Forns, 1993). En un test adaptado al individuo cada nuevo ítem por aplicar depende del éxito o fracaso del sujeto en los ítems anteriores, lo que maximiza la fiabilidad de la puntuación total y minimiza el número de ítem que se precisa aplicar para obtener una medida suficientemente fiable (Carroll, 1990). Además, estos tests simplifican el tiempo de aplicación y permiten una medida más precisa y concreta de la variable que se desee medir.

El proceso de funcionamiento informatizado de los TAC, según Lord y Stocking (1990) comprende los siguientes pasos: (1) obtener la estimación de la habilidad del sujeto; (2) seleccionar de un pool de ítems calibrados el ítem que mejor mida la habilidad en cuestión; (3) administrar y puntuar el ítem revisando la estimación del rasgo; y (4) si la estimación es suficientemente precisa se detiene el proceso, y en caso contrario se reinicia el proceso en el punto 2.

Algunos autores como Renom (1993) constatan un incremento de TACs para el ámbito educativo revisando los elementos a tener en cuenta en la construcción de un TAC. Desde su punto de vista en el proceso de construcción de un TAC en educación, se distinguen 7 facetas básicas para el desarrollo de un sistema (Renom, 1993): (1) Adecuación del contenido; (2) Condiciones del sistema Hardware-Software; (3) Calibración del banco de ítems: estimación de parámetros, dimensionalidad de los ítems, proceso de equiparación y anclaje; (4) Estrategias de administración de ítems; (5) Validez; (6) Fiabilidad (precisión en la estimación final de la capacidad); y (7) Control de factores humanos y situacionales.

Navas (1994, 2001) enumera entre las ventajas del uso de la TRI en el TAC frente al uso de los métodos de TC: (1) Permite obtener mayor información sobre el sujeto por ítem aplicado, lo que se traduce en un incremento en la precisión de la medida y en ahorro de tiempo; (2) Permite obtener una mayor validez aparente, ya que el sujeto no tiene que responder a ítems demasiado fáciles; y (3) la habilidad del sujeto y la dificultad de los ítems están en la misma escala (o métrica), lo que proporciona una pauta clara en la selección óptima de los ítems: aplicar ítems de dificultad similar al nivel estimado para la habilidad del sujeto para obtener estimaciones precisas de las mismas.

Las ventajas citadas y la disponibilidad creciente de potentes microordenadores y terminales, el software adecuado y bancos de ítems, está potenciando cada vez más el uso de este tipo de tests, con notorias ventajas sobre los clásicos. Ello no quiere decir que todos los

problemas estén resueltos, como es el caso de la correcta equiparación y de la selección de ítems o problemas específicos de estimación de los parámetros. No obstante, podemos predecir que los próximos años verán un gran crecimiento de los tests a medida y de las evaluaciones individualizadas a través de terminales y PCs.

6.2.6.3. Tests Referidos al Criterio

El caso de los Tests Referidos al Criterio (TRC), es parecido al de los Tests a Medida que, aunque de uso conocido en la psicometría, recibirán un nuevo impulso en el marco de la TRI, el cual permitirá resolver satisfactoriamente algunos de los principales problemas que tenían planteados (Muñiz y Hambleton, 1992; Navas, 1994, 2001). Como es bien sabido, se entiende por TRC aquellos tests cuyo interés, al contrario que los normativos, se centra en la determinación del grado en el que los sujetos dominan ciertas materias o áreas de conocimiento, a las que se suele denominar criterio o dominio.

La TRI viene a solucionar dos problemas básicos de los TRC, el primero, la unificación de la métrica para los distintos tests empleados (invarianza), y el segundo, la confección del test más adecuado para maximizar la información que aporta en torno al punto de corte entre los sujetos a clasificar como conocedores del área y los no conocedores.

Dado su objetivo centrado en el dominio en vez de en las normas, los TRC se prestan más a ser utilizados en áreas donde éstos aparecen claramente definidos, como el campo educativo y profesional, más que el relativo a variables psicológicas como tales. La proliferación de bancos de ítems agrupados por especialidades académico-profesionales ha supuesto un empuje para esta forma de medida, y es cada día más frecuente ver expresadas las puntuaciones de los sujetos referidas al porcentaje de materia que dominan, más que en relación a sus compañeros (tests normativos).

Como subrayan Muñiz y Hambleton (1992) un problema claro es que ese tipo de variables de rendimiento es poco plausible considerarlas unidimensionales, por lo que la generalización de los modelos multidimensionales de TRI supondrá una buena noticia para los TRC.

6.2.6.4. Equiparación de puntuaciones

El proceso de equiparación de puntuaciones es un proceso fundamental cuando se trabaja con distintos instrumentos de medida, ya que representa el medio básico de que se dispone para poder garantizar la adecuada comparación de las puntuaciones obtenidas en distintas pruebas: es el proceso que permite el desarrollo de una conversión del sistema de unidades de un test al sistema de unidades de otro.

El problema de la equiparación de las puntuaciones de los tests queda resuelto en el marco de la TRI puesto que al ser las mediciones invariables con respecto al instrumento, es decir, al ser la métrica común, no es necesario. No obstante, hay que subrayar que esto es así siempre que el banco de ítems, del cual se extraen los distintos test utilizados, esté calibrado en la misma métrica (Muñiz y Hambleton, 1992).

La gran ventaja que presenta trabajar con modelos de la TRI es que, dentro de este marco, se obvia la necesidad de equiparar las puntuaciones de los tests, por la propiedad de invarianza de los parámetros del modelo, razón por la cual algunos autores prefieren hablar de escalamiento para lograr comparabilidad en vez de equiparación (Navas, 1994).

Cuando por razones prácticas interesa hacer corresponder la puntuación obtenida en un test con la obtenida en otro, la curva característica del ítem será el instrumento de equiparación automática, puesto que asignará al mismo grado de habilidad evaluado los correspondientes valores en puntuaciones del test (Hambleton, 1993; Woodcock y Muñoz, 1993).

6.2.6.5. Sesgo

En psicometría se considera que un test (o ítem) está sesgado contra un grupo o clase de sujetos cuando este grupo o clase, a pesar de tener el mismo nivel que otros en la variable medida, sale sistemáticamente perjudicado al utilizar dicho test. El interrogante que se plantea es cómo saber que dos sujetos o grupos tienen el mismo nivel en la variable medida, pues para saberlo hay que aplicar el test, y si está sesgado no hay manera de hacer inferencias cabales acerca del nivel de los sujetos. Esta circularidad no está bien resuelta con las técnicas clásicas empleadas para detectar el sesgo de los ítems, ya que generalmente representan los aciertos a cada ítem de los grupos a evaluar frente a puntuaciones globales en el test y analizan las diferencias entre ambos, mediante alguna técnica estadística (Muñiz y Hambleton, 1992).

Bajo el enfoque de la TRI un ítem estará sesgado si su Curva Característica del Ítem (CCI) es distinta para los grupos evaluados (Navas, 1994, 2001). Por lo tanto, estiman las CCI para ambos (o más) grupos y ven en qué medida difieren. Se han propuesto varios métodos para estimar la significación estadística de la diferencia entre las CCIs, desde la mera inspección visual hasta complejos análisis estadísticos. La solución que aporta la TRI es adecuada una vez superado el problema técnico de encontrar un buen indicador de la significación estadística de las diferencias entre los CCIs en tamaños muestrales grandes.

Si bien los métodos basados en la TRI son métodos muy adecuados para la detección del sesgo, no obstante, no están exentos de problemas. Una de las cuestiones más espinosas es la relativa a los grandes tamaños muestrales necesarios para garantizar la adecuación en el funcionamiento de estos métodos. Por otro lado, requiere procedimientos sofisticados y costosos desde el punto de vista computacional, y además puede suceder que las diferencias observadas en las CCIs de los grupos se deban no a la existencia de sesgo en el ítem en cuestión sino a que el error de estimación cometido en uno y otro grupo es, sencillamente, distinto.

En síntesis, los modelos de Teoría de Respuesta al Ítem han supuesto un gran avance los últimos años en la medición psicológica y educativa. No son la panacea, ni todos sus problemas están totalmente resueltos, pero han introducido una nueva tecnología de la medida con claras aplicaciones en la práctica; de hecho muchos de los tests elaborados en la actualidad están contruidos utilizando la TRI. En un futuro el uso de estos modelos se verá incrementado a medida que se extiendan los bancos de ítems. No obstante, son varios los problemas a resolver, por ejemplo, la elaboración de modelos unidimensionales para ítems de respuesta no sólo dicotómica, sino también politómica, a la vez que modelos multidimensionales para ambos tipos de respuesta, o, los efectos sobre el ajuste de los modelos de la violación del supuesto de unidimensionalidad. En esta última dirección se han realizado algunos trabajos (Cuesta y Muñiz, 1995) que estudian los efectos de la mutidimensionalidad en la estimación de parámetros desde modelos unidimensionales de Teoría de Respuesta al Ítem. Estos investigadores concluyen que el modelo es robusto a violaciones moderadas del supuesto de unidimensionalidad (Cuesta y Muñiz, 1995).

La TRI es el avance indiscutible de la psicometría de los últimos años, trascendiendo la aportación teórica y con una incidencia en la práctica de la Evaluación Psicológica (Forns, 1993; Urbina, 2007). Las principales aportaciones se sitúan en el dominio de la evaluación informatizada, donde se dispone de grandes bancos de ítems y de posibilidad y facilidad para su adaptación a la medida del sujeto, imponiéndose más firmemente en áreas de dominios bien definidos, como pueden ser los ámbitos educativos y laborales, más que en el ámbito clínico. Las medidas clínicas, individuales, basadas en los presupuestos de la TRI, avanzan lentamente, ya que sus variables son difícilmente operacionalizables.

A pesar de sus ventajas en la investigación psicométrica, la Teoría de Respuesta al Ítem (TRI) no ha logrado imponerse en la práctica cotidiana de medición de constructos psicológicos o actitudes sociales. En la reciente investigación de Asún y Zuñiga (2008) se ha mostrado la utilidad de trabajar con modelos de TRI politómicos a través de su comparación con la Teoría Clásica, al estudiar el comportamiento de una escala de intolerancia. El estudio concluye que las ventajas de los modelos politómicos de la TRI no se encuentran en la forma en que escalan a los individuos ni en la estimación de los parámetros de los ítems, sino en la obtención de mayor información respecto al funcionamiento del instrumento, que podría ser utilizado para su mejoramiento futuro.

6.3. GARANTÍAS ÉTICAS EN LA EVALUACIÓN PSICOLÓGICA

6.3.1. EL CÓDIGO DEONTOLÓGICO DEL PSICÓLOGO

La Asociación Norteamericana de Psicología (APA) fundada en 1892 funcionó durante más de 40 años encarando problemas éticos a medida que aparecían y resolviéndolos de manera informal. En 1938, nombra su primera comisión encargada de establecer normas éticas para sus miembros y resuelve utilizar una metodología empírica y participativa sin antecedentes en otras asociaciones. Partiendo de la base de que el código debía ser representativo de las cuestiones éticas que los psicólogos enfrentan en su práctica diaria, la comisión lanza una encuesta en la que solicita ejemplos de los tipos de problemas que sus miembros encuentran conflictivos. Sobre esta metodología de incidentes críticos, su análisis y propuestas de resolución de los mismos, se sustentó la versión "*Principios éticos para psicólogos*", publicada en 1951. Los principios generales del Código de la ética incluyen: la competencia, la integridad y la responsabilidad profesional y científica del psicólogo, el respeto por los derechos y la dignidad humanos, la preocupación por el bienestar de los otros, y una activa responsabilidad social en la comunidad (Frank de Verthelyi, 1996).

La evaluación psicológica, además de unos requisitos científicos, debe de estar sujeta a ciertos principios éticos establecidos por la comunidad de referencia (Fernández-Ballesteros y Maciá, 1992). El Colegio Oficial de Psicólogos (1984, 1993), en colaboración con una serie de instituciones y profesionales entre las que se encontró la Sociedad Española de Evaluación Psicológica (SEEP), elaboró el Código Deontológico del Psicólogo, algunos de cuyos artículos hacen referencia a la Evaluación Psicológica.

Este Código Deontológico tiene por objetivo *regular el comportamiento profesional del psicólogo*: "El código deontológico de la profesión de psicólogo/a está destinado a servir como regla de conducta profesional en el ejercicio de la Psicología en cualquiera de sus modalidades. El Colegio Oficial de Psicólogos lo hace suyo y, de acuerdo con sus normas, juzgará el ejercicio de la profesión de los colegiados".

Algunas de sus normas hacen referencia a una serie de **principios generales de actuación**: “La profesión de psicólogo/a se rige por principios comunes a toda deontología profesional: respeto por la persona, protección de los derechos humanos, sentido de responsabilidad, sinceridad para con los clientes, prudencia en la aplicación de instrumentos y técnicas, competencia profesional, solidez de la fundamentación objetiva y científica de sus intervenciones profesionales”.

Por otro lado, este código regula la **formación del psicólogo** en su actuación profesional, así como el uso exclusivo y confidencial del material psicológico: “La autoridad profesional del psicólogo/a se fundamenta en su capacitación y cualificación para las tareas que desempeña. El/la psicólogo/a ha de estar profesionalmente preparado y especializado en la utilización de métodos, instrumentos, técnicas y procedimientos que adopte en su trabajo. Forma parte de su trabajo el esfuerzo continuado de actualización de su competencia profesional. Debe reconocer los límites de su competencia y las limitaciones de sus técnicas”...“Todo tipo de material estrictamente psicológico, tanto de evaluación cuanto de intervención o tratamiento, queda reservado al uso de los/las psicólogos/as, quienes se abstendrán de facilitarlos a otras personas no competentes”.

Otro grupo de artículos hacen referencia a la **utilización correcta de los conocimientos psicológicos**: “El/la psicólogo debe rechazar llevar a cabo la prestación de sus servicios cuando haya certeza de que pueden ser mal utilizados en contra de los legítimos intereses de las personas, los grupos, las instituciones, las comunidades...” . A medida que los psicólogos han ampliado sus ámbitos de intervención también se ha acrecentado el margen de la sospecha de que su saber pueda ser utilizado en contra de los derechos del individuo.

Por otro lado, este código deontológico regula el **rigor científico en la actuación profesional**: “Sin perjuicio de la legítima diversidad de teorías, escuelas y métodos, el/la psicólogo/a no utilizará medios o procedimientos que no se hallen suficientemente contrastados, dentro de los límites del conocimiento científico vigente. En el caso de investigaciones para poner a prueba técnicas o instrumentos nuevos, todavía no contrastados, lo hará saber así a sus clientes antes de su utilización”.

Así mismo se regulan el **uso del material psicológico y los datos** empleados en toda evaluación psicológica: “Los registros escritos y electrónicos de los datos psicológicos, entrevistas y resultados de pruebas, si son conservados durante cierto tiempo, lo serán bajo la responsabilidad personal del psicólogo en condiciones de seguridad y secreto que impidan que personas ajenas puedan tener acceso a ellos”.

Finalmente, otra serie de artículos agrupan una serie de **normas referentes al informe psicológico y a sus características de confidencialidad y rigor**: “Los informes psicológicos realizados a petición de instituciones u organizaciones en general, aparte de lo indicado en el artículo anterior, estarán sometidos al mismo deber y derecho general de confidencialidad antes establecido, quedando tanto el psicólogo/a como la correspondiente instancia solicitante obligados a no darles difusión fuera del estricto marco para el que fueron recabados”... “Los informes psicológicos habrán de ser claros, precisos y rigurosos para su destinatario. Deberán expresar su alcance y limitaciones, el grado de certidumbre que, acerca de sus varios contenidos, posea el informante, su carácter actual o temporal, las técnicas utilizadas para su elaboración, haciendo constar, en todo caso, los datos del profesional que lo emite”.

Dentro del código deontológico, los principios éticos de la evaluación psicológica enunciados por APA (1992) se concretan en 10 artículos.

1. Evaluación, Diagnóstico e Intervenciones en un contexto profesional

- a. Los psicólogos únicamente llevan a cabo evaluaciones, diagnósticos o intervenciones dentro del contexto de una relación profesional definida.
- b. Las evaluaciones de los psicólogos, recomendaciones, informes y diagnósticos o apreciaciones valorativas, se basan en información y técnicas (incluyendo entrevistas personales cuando sea conveniente) suficientes para proporcionar una fundamentación adecuada de sus hallazgos.

2. Competencia y uso apropiado de evaluaciones e intervenciones

- a. Los psicólogos que elaboran, aplican, puntúan, interpretan, o usan técnicas de evaluación psicológica, entrevistas, tests, u otros instrumentos, lo hacen de modo que tanto el proceder como los fines perseguidos son apropiados de acuerdo con la investigación y los datos existentes acerca de la utilidad y pertinencia de las técnicas.
- b. Los psicólogos se abstienen del uso inapropiado de técnicas de evaluación, intervenciones, resultados e interpretaciones, y toman medidas razonables para evitar que otros usen inadecuadamente la información proporcionada por estas técnicas. Esto incluye el abstenerse de facilitar las puntuaciones de los tests u otros datos a personas, ajenas a los pacientes o clientes, no cualificadas para utilizar dicha información.

3. Construcción de tests

Los psicólogos que construyen y realizan investigación con tests y otras técnicas de evaluación, utilizan procedimientos científicos y conocimientos profesionales actualizados para el diseño de los tests, estandarización, validación, reducción o eliminación del sesgo, y recomendaciones de uso.

4. Uso de la evaluación en general y con poblaciones especiales

- a. Los psicólogos que llevan a cabo intervenciones o aplican, puntúan, interpretan, o utilizan técnicas de evaluación, conocen la fiabilidad, validez, los estudios sobre la estandarización y los resultados, aplicaciones adecuadas y usos de las técnicas que manejan.
- b. Los psicólogos son conscientes de los límites de la certidumbre con la que pueden hacerse los diagnósticos, juicios y predicciones acerca de las personas.
- c. Los psicólogos intentan identificar las situaciones en las cuales determinadas intervenciones, técnicas de evaluación, o normativas pueden no ser aplicables, o pueden requerir un ajuste en la aplicación o interpretación, debido a características de la persona tales como el género, edad, raza, etnia, nacionalidad, religión, orientación sexual, discapacidad, idioma o estatus socioeconómico.

5. Interpretación de los resultados de la evaluación

Al interpretar los resultados de la evaluación, incluidas las interpretaciones automatizadas, los psicólogos tienen en cuenta los distintos aspectos del test y características de la persona evaluada que pueden afectar los juicios del psicólogo o reducir la precisión de sus

interpretaciones. Han de indicar cualquier reserva significativa que tengan acerca de la precisión o limitaciones de sus interpretaciones.

6. Personas no cualificadas

Los psicólogos no promueven el uso de técnicas de evaluación psicológica por parte de personas no cualificadas.

7. Tests obsoletos y resultados desfasados de los tests

- a. Los psicólogos no basan sus evaluaciones, decisiones sobre la intervención, o recomendaciones, en datos o resultados de los tests que están desfasados en relación con los objetivos perseguidos en la actualidad.
- b. Del mismo modo, los psicólogos no basan sus decisiones o recomendaciones en medidas y tests obsoletos, inapropiados para los objetivos actuales.

8. Servicios de calificación e interpretación

- a. Los psicólogos que ofrecen procedimientos de calificación o evaluación a otros profesionales describen con precisión la finalidad, normas, validez, fiabilidad y aplicaciones de los procedimientos, así como cualquier cualificación especial aplicable a su uso.
- b. Los psicólogos eligen los servicios de calificación e interpretación (incluidos los servicios automatizados) basándose en los datos sobre la validez del programa y procedimientos, así como en otras consideraciones pertinentes.
- c. Los psicólogos mantienen la responsabilidad de la aplicación adecuada, la interpretación y el uso de los instrumentos de evaluación, tanto si son ellos mismos los que puntúan e interpretan los tests, como si utilizan servicios automatizados u otros para hacerlo.

9. Explicación de los resultados de la evaluación

Al menos que se informe claramente con antelación a la persona evaluada que la evaluación excluye cualquier tipo de explicación de los resultados (como ocurre con algunas situaciones de consejo organizacional, o ciertas evaluaciones de preempleo, seguridad, o del ámbito jurídico), los psicólogos han de garantizar que se proporcione una explicación de los resultados, utilizando un lenguaje que pueda entender la persona evaluada, u otra persona legalmente autorizada por el cliente. Independientemente de que la puntuación e interpretación sean hechas por el propio psicólogo, por ayudantes, o por medio de servicios automatizados u otros servicios externos, los psicólogos toman las medidas necesarias para asegurar que se proporcionan explicaciones apropiadas de los resultados.

10. Mantenimiento de la seguridad de los tests

Los psicólogos se esfuerzan para mantener la integridad y seguridad de los tests y otras técnicas de evaluación, de acuerdo con la ley y las obligaciones contractuales, y de tal modo que le permita cumplir con los requerimientos del código ético.

Teniendo en cuenta los códigos deontológicos y las normas éticas de los psicólogos de la *American Psychological Association* (APA 1992), Sattler (1996) plantea varias cuestiones referidas a la evaluación:

- Los psicólogos deben conocer que sus recomendaciones pueden alterar significativamente la vida de los demás y han de asumir, por lo tanto, sus responsabilidades sociales.
- Los psicólogos deben reconocer sus propias competencias y las limitaciones anejas a ellas y a la instrumentación que les es propia.
- Los psicólogos deben proteger la confidencialidad de la información obtenida en el curso de la evaluación.
- Los clientes tienen derecho a acceder a los resultados de la evaluación y se les debe explicar las bases que han tenido las conclusiones y recomendaciones que se han hecho.
- Los tests y sus resultados deben guardarse debidamente.
- Si existen dudas respecto a la fiabilidad o validez de los resultados de los tests, éstas deben quedar claramente expresadas en el informe.

En el ámbito específico de la evaluación psicológica con niños y adolescentes, recientemente Silva (1995c) focaliza su atención sobre 5 cuestiones deontológicas:

- Justificación de la evaluación: una evaluación debe estar justificada, porque de lo contrario sería una intromisión injustificada en el ámbito privado de las personas.
- Consentimiento: Tres condiciones son necesarias para que un consentimiento sea tanto ética como legalmente aceptable, competencia o capacidad de consentir, que sea voluntario y que esté informado.
- Confidencialidad: Derecho a la intimidad.
- Base psicométrica deficiente o insuficiente: Utilizaciones de tests para distintos fines de aquellos para los que fueron contruidos y validados; Violación de las normas de aplicación de los instrumentos;....
- Cuidado del vínculo interpersonal: Preparar al niño para entrar en la relación interpersonal con el psicólogo y para salir de ella.... informando al niño de qué trata la relación que comienza, cómo transcurrirá y qué se espera de las personas participantes, en especial qué puede él esperar y no esperar del psicólogo.

La ética profesional nos obliga a reconocer que nuestra tarea afecta la vida de los otros, para bien o para mal, y que somos responsables de nuestras acciones o inacciones, de nuestra buena o mala formación. Los códigos y los principios éticos pueden indicarnos el camino correcto y sancionarnos si nos desviamos de él. Pero no deben ser sustitutos de una constante revisión personal respecto de nuestros valores en relación con el desempeño del rol. Ejercerlo con ética es siempre nuestra responsabilidad individual.

Los *Principios éticos y el Código de conducta de los psicólogos* elaborado por la *American Psychological Association* (APA, 2002) recomienda que antes de la aplicación de los tests psicológicos y otros procedimientos de evaluación los psicólogos obtengan un *consentimiento informado de los sujetos*, bien oralmente o por escrito. Para que el consentimiento pueda considerarse informado debe estar precedido de una adecuada explicación de la naturaleza y la

finalidad de la evaluación, y de una información sobre los límites de confidencialidad y la forma en que se mantendrá la seguridad de los resultados de los tests. En ciertos casos concretos deberán comentarse otros aspectos prácticos que puedan afectar a los sujetos (por ejemplo, coste de la prueba, posibilidad de rechazar o abandonar la prueba, terceras personas implicadas en el proceso...). Los sujetos deben tener la oportunidad de formular preguntas sobre el proceso de evaluación, que deben constestarse adecuadamente (Urbina, 2007). El código de ética de los psicólogos (APA, 2002) enumera algunos casos excepcionales en los que no se requiere el consentimiento informado, bien porque se da por supuesto o bien porque la evaluación es obligada, mencionando las condiciones especiales que se dan en estos casos y las personas que no tienen capacidad de dar su consentimiento.

La revisión de los códigos deontológicos en diferentes países realizada por Fernández-Ballesteros y Calero (2004) les lleva a concluir que como novedad, con respecto a los estándares americanos, que les sirvieron de base, el código deontológico español, introduce un artículo mediante el cual recomienda prudencia en la utilización de etiqueta negativa y desarrolla pormenorizadamente los artículos referidos a la propiedad de la información recogida en el acto evaluativo. Podemos decir que nuestro código deontológico se dirige fundamentalmente a regular la psicología aplicada, distinguiéndose de la actual normativa americana, más interesada en la protección de los derechos de los sujetos de investigación sean humanos y/o animales.

En este sentido el código deontológico, actualmente vigente en nuestro país, se dirige a proteger la práctica profesional mediante la utilización de instrumentos contrastados, sobre los cuales se debe informar al sujeto o a sus padres o tutores, que den lugar a informes psicológicos claros, rigurosos e inteligibles en los que se recojan no sólo las técnicas utilizadas, sino también el grado de certidumbre de los resultados de las mismas. Se dirige además, a proteger el derecho a la intimidad, y al secreto profesional que tiene el individuo (institucionalizado o no), dueño de la información sobre sí mismo junto con aquellos que hayan solicitado el informe, sean allegados o instituciones. Curiosamente, otra diferencia de nuestro código con respecto al americano es el interés de este segundo por los aspectos relacionados con la incorporación del ordenador a la psicología aplicada (Fernández-Ballesteros y Calero, 2004). Progresivamente, en España, la preocupación por la vulneración del código deontológico está cobrando mayor relieve y preocupación, lo que se pone de relieve tanto en la publicación de libros sobre el tema (Del Río, 2005) como en artículos en los que se aborda esta temática (Bermejo, 2007; Del Río, 2007).

Para finalizar, dado que nuestra sociedad cada vez es más plural, es necesario hacer algunas referencias sobre la necesidad de una *especialización multicultural*. Recogiendo algunas de las observaciones de Kamphaus y Frick (2000) se exponen las consideraciones con respecto a la evaluación de niños y adolescentes que difieren del evaluador en variables culturales. Al trabajar con otras poblaciones cultural o socialmente diferentes, el profesional para poder realizar una evaluación competente debe conocer bien la historia, cultura y lengua de dichos grupos a fin de no emplear procedimientos inadecuados e realizar interpretaciones sesgadas. Geisinger (1992) recomienda una serie de pautas en la evaluación de sujetos culturalmente diferentes:

- Conocer la investigación pertinente y aspectos prácticos relacionados con los grupos a los que se atiende.

- Reconocer que el origen étnico y cultural son parámetros significativos para la comprensión de los procesos psicológicos.
- Ayudar a los clientes a incrementar el reconocimiento de sus propios valores y normas culturales, y facilitar el descubrir modos por los cuales los clientes puedan aplicar este reconocimiento a sus propias vidas y a la sociedad en su conjunto.
- Respetar los roles de los miembros de la familia, así como las estructuras, jerarquías, valores y creencias de la cultura del cliente.
- Respetar los valores, creencias religiosas y espirituales del cliente, incluyendo atributos y tabúes, ya que afectan la visión global, el funcionamiento psicosocial y las expresiones de angustia.
- Interactuar en el idioma solicitado por el cliente y, si esto no es posible, ofertar servicios de un traductor con conocimiento cultural, y adecuados antecedentes profesionales.
- Considerar el impacto de los factores adversos de tipo sociopolítico y del ambiente al evaluar problemas y diseñar intervenciones.
- Atender con el fin de eliminar prejuicios, sesgos y prácticas discriminatorias.
- Documentar en los registros e informes los factores culturales y sociopolíticos que se consideren relevantes.

Estas pautas relacionadas con la evaluación de los aspectos de la diversidad sugieren de manera indirecta la necesidad de evaluar con mayor precisión el nivel de aculturación del sujeto.

En los últimos tiempos se está desarrollando un movimiento que favorece no sólo la aparición de normas y guías, sino también un proceso de unificación de criterios entre países, y eso se ha ido plasmando en una serie de documentos de gran interés para el evaluador y el psicólogo aplicado (Fernández-Ballesteros y Calero, 2004).

6.3.2. DIRECTRICES ASOCIADAS AL USO DE LOS TESTS

6.3.2.1. Directrices internacionales para el uso de los tests

Los tests constituyen una de las tecnologías más utilizadas por los psicólogos en el ejercicio de la profesión. Como ocurre con la tecnología de otras áreas científicas, los tests pueden utilizarse correcta o incorrectamente. Las directrices que propone la comisión Internacional de Tests (ITC) tienen como objetivo fundamental mejorar el uso que los profesionales hacen de los tests.

Las directrices han sido elaboradas originalmente por la *Comisión Internacional de Tests (ITC)*, bajo la dirección del profesor Dave Bartram. La versión en castellano que se presenta aquí ha sido realizada en el seno de la *Comisión de Tests* del Colegio Oficial de Psicólogos (COP). El Colegio, a través de sus representantes en la citada Comisión Internacional de Tests y en la Comisión Europea sobre Tests de la Federación Europea de Asociaciones Profesionales de Psicólogos (EFPPA), que también ha respaldado las directrices, ha participado activamente en la elaboración de las mismas, consciente de la necesidad de contribuir a mejorar la práctica de los tests en nuestro país.

Las directrices reflejan principios generales implicados en el uso adecuado de los tests, pero no pretenden uniformar las diferencias legítimas existentes entre la práctica en los

diferentes países o áreas profesionales. Existen bastantes razones por las que se necesitan unas directrices para el uso de los tests con carácter internacional. Veamos algunas de ellas.

- a) Los países difieren ampliamente en el grado de control legal que pueden ejercer sobre el uso de los tests y sus consecuencias para las personas evaluadas. Por tanto, la existencia de un conjunto de directrices internacionalmente aceptadas proporcionará a las asociaciones nacionales de psicología y otras organizaciones profesionales una buena base documental para completar directrices ya existentes o proceder a crearlas.
- b) El derecho a comprar y utilizar materiales psicométricos varía mucho de unos países a otros. En algunos países el uso de los tests está restringido a los psicólogos, en otros a quienes estén registrados por los distribuidores de las pruebas, e incluso en algunos los usuarios pueden obtener el material psicométrico libremente, sin restricciones.
- c) Algunos tests bien conocidos han aparecido en Internet sin la autorización de autores y editores, violando los derechos de *copyright*, y sin ninguna precaución en relación con la seguridad del test.
- d) En el ámbito de la psicología del trabajo, el aumento de la movilidad internacional de los trabajadores ha incrementado la demanda de tests para utilizarlos con aspirantes de diferentes países. A menudo los tests son aplicados en un país para empresas ubicadas en otros países.
- e) En países como Estados Unidos e Inglaterra, entre otros, se están llevando a cabo trabajos sobre evaluación a distancia vía Internet en las áreas de Trabajo y Educativa. Esto plantea todo un conjunto de problemas en relación con las normas de aplicación, control del proceso de evaluación y seguridad de los resultados.

Por todas estas razones, y otras que se podrían enumerar, parece conveniente disponer de unas directrices internacionales que permita una referencia común a la hora de establecer normas nacionales para la mejora del uso de los tests.

Las directrices de la ITC recogen el trabajo de especialistas en evaluación psicológica y educativa (psicólogos, psicómetras, editores y constructores de tests) de diferentes países, entre ellos España. La idea no ha sido tanto el inventar unas nuevas directrices, como juntar las hebras comunes que se encuentran en las distintas directrices, códigos de práctica y estándares ya existentes, formando una estructura coherente que resulte comprensible y de uso práctico.

Objetivo: El objetivo fundamental de las directrices es mejorar el uso de los tests, describiendo la forma adecuada de utilizarlos. Un usuario competente utilizará los tests de forma adecuada, profesional y ética, prestando la debida atención a las necesidades y derechos de las personas implicadas en el proceso de evaluación, y teniendo muy en cuenta las razones para utilizar los tests, así como el contexto en el cual se lleva a cabo su aplicación. Este objetivo se alcanzará asegurándose de que el usuario de los tests tiene las competencias y conocimientos necesarios para llevar a cabo el proceso evaluativo.

Ámbito de aplicación: Resulta difícil dar una definición muy precisa de *test* o de su *práctica*, pues al hacerlo pueden dejarse fuera algunos procedimientos que debieran incluirse, o, por el

contrario, incluir otros que deberían quedar fuera. En las directrices de la ITC los términos *test* y *práctica de los tests* se utilizan en sentido amplio. Las directrices son aplicables a muchos procedimientos de evaluación que no suelen denominarse *tests*, o que incluso tratan de evitar su designación como *tests*. Más que proponer una definición cerrada, se describe a continuación el ámbito que tratan de cubrir las directrices.

- Los tests incluyen un amplio abanico de procedimientos utilizados en la evaluación psicológica, educativa y ocupacional.
- Los tests incluyen procedimientos para la medición de conductas tanto normales como anormales o disfuncionales.
- Los tests son procedimientos diseñados para ser aplicados bajo condiciones controladas o estandarizadas, y conllevan la utilización de protocolos de puntuación rigurosos.
- Estos procedimientos proporcionan medidas de ejecuciones y conllevan la obtención de inferencias a partir de muestras de conducta. También pueden incluir procedimientos que proporcionan clasificaciones cualitativas u ordenamientos de las personas.

Los tests deben de estar apoyados por datos empíricos sobre su fiabilidad y validez para medir los objetivos que se proponen. Hay que aportar datos que justifiquen las inferencias que se hacen a partir de las puntuaciones de los tests. Estos datos tienen que estar disponibles para los usuarios de los tests, así como para los profesionales e investigadores que deseen llevar a cabo una evaluación o revisión independiente.

Las directrices deben de tenerse en cuenta para cualquier procedimiento de evaluación utilizado en situaciones en las que la evaluación tiene serias implicaciones para las personas, pudiendo causarles daños personales o psicológicos si no se realiza adecuadamente.

Aspectos contextuales: Las directrices son aplicables internacionalmente, pero a la hora de utilizarlas a nivel local en un país determinado deben de tenerse en cuenta diversos factores o aspectos que pueden modular su aplicación. Entre otros, estos factores contextuales incluyen:

- Diferencias sociales, políticas, institucionales, lingüísticas y culturales.
- La legislación del país en el que se utilizan los tests.
- Directrices y estándares ya existentes establecidos por asociaciones profesionales.
- Diferencias entre la evaluación individual y la colectiva.
- Diferencias en función de la situación evaluativa (educativa, clínica, trabajo...).
- Quienes son los receptores de los resultados (los propios evaluados, los padres o responsables legales, los constructores de los tests, los empleadores, etc.)
- Diferencias según el uso de los resultados de los tests (por ejemplo, toma de decisiones frente a orientación).
- Grado en el que la situación permite la posibilidad de comprobar la precisión de las predicciones y su posible modificación a la luz de los resultados posteriores.

Conocimientos: Unos conocimientos psicológicos psicométricos sólidos y una comprensión profunda de todos los aspectos implicados en el proceso evaluativo constituyen la base fundamental para el uso pertinente de los tests. Los expertos suelen estar de acuerdo en que la causa más importante del uso inapropiado de los tests es una formación deficiente de los

usuarios. Los conocimientos concretos requeridos en cada caso variarán en función de la situación y área de aplicación, por tanto no es fácil describir de forma general los conocimientos requeridos para la utilización adecuada de los tests en todas las situaciones, depende de cada caso. Por ejemplo, no se requieren los mismos conocimientos para aplicar un test colectivo a un grupo que para interpretar las puntuaciones, o construir el propio test, por citar un caso obvio. Las directrices no hacen descripciones detalladas de estos conocimientos, ni lo pretenden. No obstante, a la hora de aplicarlas a una situación concreta habrá que tener muy presente y describir con precisión los conocimientos requeridos para el caso. Esta descripción tiene que cubrir las distintas áreas de conocimientos implicadas en la situación de que se trate, entre ellas deberían de incluirse las siguientes: a) Conocimientos sobre teoría de los tests y propiedades técnicas de los tests, tales como fiabilidad, validez, estandarización, sesgo, análisis de ítems, etc; b) Conocimiento de los tests y principios de la medición para entender adecuadamente los resultados; c) Conocimientos sobre la teoría, modelos y constructos medidos, que permita una elección pertinente de las pruebas e interpretación de los resultados; d) Conocer el abanico de tests disponibles, así como los editores correspondientes; e) Avances técnicos recientes, tales como los tests informatizados, bancos de ítems, etc. Otros muchos conocimientos específicos han de ponerse en funcionamiento para llevar a cabo una utilización adecuada de los tests, cada una de las directrices puede servir como indicador de la naturaleza de éstos.

Articulado: Los 135 artículos que contienen las directrices internacionales para el uso de los tests, se dividen en dos apartados: 1) uso ético de los tests; y 2) utilización adecuada de los test.

1. Uso ético de los tests

- 1.1. Actuar de forma ética y profesional.
- 1.2. Asegurarse de que son competentes para el uso de los tests.
- 1.3. Responsabilizarse del uso que hacen de los tests.
- 1.4. Asegurarse de que los materiales del test están seguros.
- 1.5. Asegurarse de que los resultados de los tests se tratan confidencialmente.

2. Utilización adecuada de los tests

- 2.1. Estimar la utilidad potencial de los tests en una situación evaluativa.
- 2.2. Elegir tests técnicamente correctos y adecuados a cada situación.
- 2.3. Prestar atención a los aspectos relacionados con el sesgo de los tests.
- 2.4. Hacer los preparativos necesarios para la aplicación del test.
- 2.5. Aplicar los tests adecuadamente.
- 2.6. Puntuar y analizar los resultados de los tests con precisión.
- 2.7. Interpretar los resultados adecuadamente.
- 2.8. Comunicar los resultados de forma clara y precisa.
- 2.9. Revisión de la adecuación del test y de su uso.

Y además, las directrices contienen tres apéndices: 1) Directrices para el establecimiento de políticas sobre el uso de los tests; 2) Directrices para desarrollar contratos entre las partes implicadas en la evaluación; y 3) Aplicación de los tests a personas con alguna discapacidad. Toda la información sobre las directrices puede encontrarse en la página web del Colegio Oficial de Psicólogos (<http://www.cop.es/vernumero.asp?id=40>)

6.3.2.2. Directrices para la traducción y adaptación de tests

Como ya subrayaron hace más de una década, Muñiz y Hambleton (1996), los tests en sus distintas modalidades constituyen uno de los instrumentos de medida más utilizados por los profesionales de la psicología y por los investigadores. Como cualquier otro instrumento de medida en cualquier campo de actividad, su utilización incorrecta resulta nefasta, pero ello es especialmente grave en psicología y educación, donde con frecuencia están implicadas decisiones de suma importancia para la vida de las personas. La mayoría de los problemas en torno a los tests provienen de su uso inadecuado, más que del test en sí, de su construcción, o de sus propiedades técnicas. Este uso inapropiado por parte de personas sin una formación suficiente, es el responsable de muchas de las críticas que reciben los tests, cuestionarios y las escalas psicológicas. Pero la realidad es que los tests siguen siendo una instrumentación imprescindible en la práctica psicológica y educativa en todo el mundo, y todo indica que va a seguir siéndolo en el futuro, por tanto no hay que ahorrar esfuerzos en mejorar su uso. Cómo hacerlo es una cuestión abierta, que ofrece muchas posibilidades y modelos, y cualquier opinión es bienvenida.

Los Colegios de Psicólogos de la mayoría de los países tratan de mejorar esta práctica de formas muy diversas. Una línea clave de actuación es promover que los tests que se comercializan estén técnicamente bien contruidos. A la base de ello está una formación psicométrica adecuada y la publicación de normas técnicas que deben de seguir los constructores. Los estándares técnicos más utilizados son los elaborados por la *Asociación Americana de Psicología*. Las propias casas editoras de tests, primeras interesadas en un uso deontológico apropiado, así como los Colegios Profesionales, toman algunas medidas complementarias, tales como restringir el acceso a los tests en función de su especialización, registrar oficialmente a los usuarios cualificados, etc.

La creciente globalización mundial que se está produciendo a todos los niveles y en todas las esferas de la actividad humana, debido a la facilidad y rapidez de las comunicaciones, potencia que los instrumentos de medida generados en determinado país rápidamente se utilicen en otro distinto. Existe además un interés creciente en todo lo relacionado con los estudios interculturales. Ahora bien, los tests no son automáticamente utilizables, deben de adaptarse cuidadosamente en función de las diferencias interculturales entre el idioma/cultura originario y aquéllos en los que se pretende utilizar. Esto es especialmente importante en un país como España en el que un porcentaje muy elevado de los tests y las escalas que se utilizan provienen de traducciones y adaptaciones de otros idiomas y culturas.

Con la inminente integración europea el problema seguramente se agudizará. Para no sesgar de entrada el uso de estos tests hay que asegurarse de que su traducción/adaptación es la adecuada. Si ya apareciesen fallos a este nivel básico, cualquier medida posterior encaminada a mejorar su uso caería en saco roto. Como se irá viendo, traducir un test no es una mera cuestión de conocer bien los dos idiomas implicados, es un proceso mucho más

complejo, por eso más que de traducción habría que hablar de adaptación. Consciente de este problema, el Colegio Oficial de Psicólogos (COP) ha participado en un Comité Internacional auspiciado por la Comisión Internacional de Tests (ITC), de la cual es miembro, para preparar unas directrices técnicas para la adaptación de los tests de unos países a otros. Este Comité ha elaborado un conjunto de veintidós directrices o recomendaciones, en 4 áreas:

Contexto: Agrupa 2 normas que hacen referencia a la minimización de las diferencias entre poblaciones no relevantes para el estudio que se pretende y a la significación del constructo evaluado para la población objetivo, entendiendo que muchos de los constructos evaluados en psicología son dependientes de aspectos culturales y que por tanto no son equiparables entre culturas diferentes (lo que directamente exige la demostración de equivalencia cultural del constructo antes de llevar a cabo la adaptación del instrumento).

- C.1. Los efectos de las diferencias culturales que no sean relevantes para los objetivos centrales del estudio deberían minimizarse en la medida de lo posible.
- C.2. Debería de evaluarse la cuantía del solapamiento de los constructos en las poblaciones de interés.

Desarrollo o adaptación de instrumentos: Consta de 10 principios en los que se establecen las normas que deben regir en la traducción, tanto en los ítems como en las alternativas de respuesta, las exigencias que deben seguir el desarrollo de instrumentos para ser utilizados en estudios transculturales, la equivalencia de tests y de elementos y la validación para diferentes poblaciones.

- D.1. Los constructores/editores de tests deberían de asegurar que el proceso de adaptación tiene en cuenta las diferencias lingüísticas y culturales entre las poblaciones a las que se dirigen las versiones adaptadas del test.
- D.2. Los constructores/editores de los tests deberían de proporcionar datos que garanticen que el lenguaje utilizado en las instrucciones, en los propios ítems y en el manual del test, son apropiados para todas las poblaciones culturales e idiomáticas a las que va dirigido el test.
- D.3. Los constructores/editores de tests deberían de aportar evidencia de que las técnicas de evaluación elegidas, los formatos de los ítems, las reglas de los tests, y los procedimientos son familiares a todas las poblaciones a las que van dirigidos.
- D.4. Los constructores/editores de tests deberían de facilitar evidencia de que el contenido de los ítems y los materiales de los estímulos son familiares para todas las poblaciones a las que van dirigidos.
- D.5. Los constructores/editores de tests deberían aportar una justificación racional sistemática, tanto lingüística como psicológica, para mejorar la precisión del proceso de adaptación, así como reunir datos acerca de la equivalencia de todas las versiones en los distintos idiomas.

- D.6. Los constructores/editores de tests deberían asegurar que el diseño de recogida de datos permite el uso de técnicas estadísticas apropiadas para establecer la equivalencia entre los ítems correspondientes a las diferentes versiones idiomáticas del test
- D.7. Los constructores/editores de tests deberían aplicar técnicas estadísticas apropiadas para:
 - 1) establecer la equivalencia entre las diferentes versiones de un test; y 2) identificar componentes problemáticos o aspectos del test que puedan ser inadecuados para alguna de las poblaciones a las que va destinado el test.
- D.8. Los constructores/editores de tests deberían proporcionar información sobre la evaluación de la validez en todas las poblaciones objetivo a las que va dirigido el test adaptado.
- D.9. Los constructores/editores de tests deberían aportar datos estadísticos sobre la equivalencia de los tests para todas las poblaciones a las que van dirigidos.
- D.10. No deben utilizarse preguntas no equivalentes en todas las versiones dirigidas a diferentes poblaciones cuando se prepara una escala común, o cuando se comparan estas poblaciones. Sin embargo, pueden ser útiles para reforzar la validez de contenido de las puntuaciones de cada población por separado.

Administración: Formado por 6 directrices que explicitan las normas relativas a instrucciones, relación evaluador-evaluado y condiciones contextuales que afectan a la equivalencia de los tests.

- A.1. Los constructores y los aplicadores de los tests deberían tratar de prever los tipos de problemas que cabe esperar, y tomar las medidas oportunas para evitarlos mediante la preparación de materiales e instrucciones adecuados.
- A.2. Quienes aplican los tests deberían de ser sensibles a cierto número de factores relacionados con los materiales utilizados para los estímulos, los procedimientos de aplicación, y las formas de respuesta, que pueden reducir la validez de las inferencias extraídas de las puntuaciones.
- A.3. Aquellos aspectos del entorno que influyen en la aplicación del test deberían de mantenerse lo más parecidos posible para todas las poblaciones a las que va dirigido el test.
- A.4. Las instrucciones para la aplicación del test en el idioma fuente y en el objetivo deben minimizar la influencia de fuentes de variación no deseadas.
- A.5. El manual del test debería de especificar todos los aspectos del test y de su aplicación que han de revisarse al utilizarlo en un nuevo contexto cultural.
- A.6. El aplicador no debe de interferir, debiendo minimizarse su influencia sobre los examinados. Deben de seguirse al pie de la letra las reglas explícitas descritas en el manual del test.

Interpretación: Comprende 4 normas que tienen en consideración las diferencias culturales en las normas de interpretación de un test. La información sobre todo el proceso de adaptación que permita aportar datos para la interpretación de un resultado.

- 1.1. Cuando se adapta un test para utilizarlo en otra población, debe de facilitarse la documentación sobre los cambios, así como los datos acerca de la equivalencia entre las versiones.
- 1.2. Las diferencias entre las puntuaciones obtenidas por las muestras a las que se aplicó el test no deben de tomarse sin más directamente. El investigador tiene la responsabilidad de sustanciar las diferencias con otros datos empíricos.
- 1.3. Las comparaciones entre poblaciones sólo pueden hacerse al nivel de la invarianza que se haya establecido para la escala en la que se expresan las puntuaciones.
- 1.4. El constructor del test debería de proporcionar información específica acerca de las distintas formas en las que los contextos socioculturales y ecológicos de las poblaciones pueden afectar al rendimiento en el test, y debería sugerir procedimientos para tener en cuenta estos efectos en la interpretación de los resultados.

La creciente globalización que se está dando a nivel mundial, el problema de la emigración y sus necesidades de adaptación, el intercambio de estudiantes entre distintos países fortalecido, además, por la creación de un Espacio Europeo de Educación Superior (EEES), y el creciente interés que actualmente tienen muchos países por establecer unos estándares educativos internacionales que permitan comparar los progresos que, cada uno de ellos, han hecho en relación con los demás, son algunos de los factores que han potenciado la preocupación porque los instrumentos de medición utilizados sean los adecuados y cumplan los requisitos psicométricos necesarios para su utilización en estudios comparativos interculturales. Las diferencias encontradas en el rendimiento obtenido por los estudiantes de los distintos países, han de deberse a diferencias reales y no a otra serie de factores (diferencias culturales, idiomáticas, etc.) que invaliden la interpretación de los resultados.

Con esta preocupación de base, el objetivo del reciente trabajo titulado “la adaptación de los tests en estudios comparativos interculturales” de Barbero, Vila y Holgado (2008) ha sido analizar brevemente las fases principales que se deben cubrir a la hora de llevar a cabo la adaptación de un test a distintas culturas o idiomas, siguiendo las pautas marcadas por la Comisión Internacional de Tests, con el fin de conseguir esos requisitos psicométricos que antes se comentaban, y hacerlo de una manera sencilla, que pueda resultar útil a todas aquellas personas que, en alguna medida, estén relacionadas con el problema de la adaptación de instrumentos para la evaluación psicológica y/o educativa. Entre las fases del proceso de adaptación estos autores identifican 4 fases del proceso de adaptación en las que se pueden producir errores que hay que tratar de minimizar:

- 1) Contexto: Equivalencia del constructo en las distintas culturas en las que se va a aplicar el test para poder comparar los resultados obtenidos.

- 2) Construcción y adaptación: Tener en cuenta las peculiaridades de la población a la que se va a aplicar; por lo tanto, es fundamental que aquellas personas que van a llevar a cabo la adaptación, además de tener conocimientos acerca del proceso de construcción de un test, conozcan los idiomas correspondientes a las distintas poblaciones y las peculiaridades culturales de cada una de ellas.
- 3) Aplicación: Se deben cuidar al máximo aspectos tales como las instrucciones que se den antes de la aplicación y las interacciones entre el aplicador y los examinados ya que pueden influir en la fiabilidad y validez del test. Los aplicadores deberán seleccionarse entre las personas pertenecientes a la población a la que se aplica el test; de esta manera estarán familiarizados con los distintos matices propios de dicha cultura; por otra parte es conveniente que reciban un cierto entrenamiento sobre la forma de aplicación, haciéndoles ver la importancia que tiene el seguir al pie de la letra las instrucciones de aplicación.
- 4) Interpretación de las puntuaciones: Para poder interpretar adecuadamente los resultados obtenidos al aplicar un test, tanto si es un test original como cualquier adaptación, es necesario tener una cierta formación psicométrica; pero, en este último caso, existe además el peligro de que al comparar los resultados obtenidos en distintos grupos o países se intente hacer algún tipo de clasificación de los sujetos. Una cosa es analizar las semejanzas y diferencias entre los grupos y otra establecer cualquier tipo de clasificación, puesto que es prácticamente imposible encontrar dos comunidades que sean equiparables completamente. Para una mejor interpretación de los resultados, el manual del test deberá incluir una información exhaustiva de todo el proceso de adaptación.

6.3.2.3. Directrices internacionales para la evaluación informatizada y test aplicados por internet.

El gran incremento de los tests contruidos para ser aplicado por ordenador y por Internet ha provocado que en marzo de 2003, la International Tests Commision (ITC) enviara un borrador sobre estándares de aplicación internacional a los Colegios Profesionales (Bartram y Coiné, 2003). Este documento fue debatido en el International Congress of Psychology (Beijing, 2004), y en el aparecen cuatro objetivos fundamentales (Fernández-Ballesteros y Calero, 2004):

1. Cuidar las cuestiones tecnológicas

- Desarrollan consideraciones de hardware y software y centran su atención en la descripción de los requisitos del sistema, la justificación del programa y las posibilidades de actualización.
- Se dirigen a asegurar la robustez del programa y señalan la necesidad de anticipar posibles fallos y dar todo el soporte técnico a los usuarios.
- Se centran en tomar en consideración los factores humanos que puedan afectar a la presentación del material y tratan fundamentalmente del diseño, el tipo de estímulos, la distinción entre instrucciones e ítems, la facilidad del usuario para volver atrás, la familiaridad de los usuarios con las demandas del programa, la compatibilidad con diferentes sistemas informáticos...

- Tratan de considerar el ajuste de las características técnicas con diferentes causas de incapacitación en los sujetos evaluados y se centran en asegurar y explicitar quiénes pueden y no pueden ser evaluados mediante ese procedimiento.
- Se interesan en el suministro de ayuda e información práctica para los usuarios. Si las directrices del programa son claras, completas, explícitas; si proporcionan práctica inicial con el procedimiento...

2. Garantías de calidad

- Trata de asegurar el conocimiento y la competencia para los usuarios y los sujetos evaluados.
- Se centra en los requisitos psicométricos del procedimiento.
- Se refiere a los procedimientos que son versiones de tests clásicos de papel y lápiz, y se dirige a la demostración de la equivalencia de formatos.
- Se refiere a los análisis de fiabilidad de puntuaciones y resultados obtenidos en estas versiones.
- Trata sobre si la interpretación de resultados es apropiada señalando sus limitaciones. Si las especificaciones del programa suministran documentación y/o feedback adicional y señalan en que criterio, reglas o procedimientos basan la interpretación...
- Pretende asegurar la igualdad de acceso para cualquier grupo, país, étnia o comunidad lingüística.

3. Asegurar los niveles de control apropiados

- Sobre las condiciones de administración.
- Sobre las condiciones de supervisión.
- Sobre la práctica inicial del evaluador y la exposición inicial del evaluado.
- Sobre las posibilidades de autenticación y estafa del acceso al sistema.

4. Asegurar la privacidad y seguridad de los datos recogidos

- Seguridad para el material que constituye el test.
- Seguridad para la transferencia de datos por Internet.
- Confidencialidad de los resultados obtenidos.

En estos estándares se entiende que el uso fundamental de esta evaluación es la orientación y selección fundamentalmente profesional (Fernández-Ballesteros y Calero, 2004).

6.4. RESUMEN: GARANTÍAS CIENTÍFICAS Y ÉTICAS EN LA EVALUACIÓN PSICOLÓGICA

Con la finalidad de clarificar algunas de las cuestiones abordadas en este capítulo, a continuación se presentan varios cuadros que sintetizan: Las propiedades que los tests y las técnicas estándar de medida deben tener (validez y fiabilidad) (Cuadro 19), la teoría de la generalizabilidad (Cuadro 20), las aplicaciones de la teoría de respuesta al ítem (Cuadro 21) y algunas puntos relevantes del comportamiento ético de la figura del psicólogo (Cuadro 22).

Cuadro 19. Propiedades de los tests y de las técnicas estándar de medida

PROPIEDADES DE LOS TESTS Y DE LAS TÉCNICAS ESTÁNDAR DE MEDIDA	
FIABILIDAD Consistencia y Precisión Precisión o grado de exactitud de la medida	Fiabilidad Test-Retest Medidas repetidas del mismo fenómeno producen informaciones equivalentes. El intervalo de tiempo en las dos aplicaciones del test no debe ser superior a los 6 meses. Consistencia Interna En qué medida los elementos que componen la prueba son homogéneos o consistentes entre sí. Fiabilidad interjueces o del evaluador Grado en el cual lo obtenido por un observador es semejante a lo obtenido por otros.
VALIDEZ Definición clásica “Grado en que el instrumento mide en realidad lo que pretende medir”. Definición estándares AERA, APA, NCME, (1999) “Grado en que las pruebas acumuladas apoyan que la interpretación propuesta a partir de las puntuaciones del test es adecuada para la finalidad pretendida”	DEFINICIONES CLÁSICAS Validez de Contenido (muestral) Grado en que los ítems de un instrumento comprenden una muestra representativa de la conducta que se pretende medir. Validez Criterial (empírica) Grado en que las puntuaciones de una variable (predictora) sirven para inferir el rendimiento en otra variable (criterio). El criterio puede ser cualquier conducta que se desee predecir a partir de las puntuaciones observadas en el test. Se distinguen dos formas según el tiempo que existe entre la aplicación del test y la medición del criterio: • Validez concurrente: Validez criterial que es estimada obteniendo puntuaciones al mismo tiempo (o en tiempos muy próximos) en el test y en el criterio. El coeficiente de validez es el coeficiente de correlación entre las puntuaciones obtenidas en el test y el criterio. Un test con validez concurrente permite contestar preguntas como ¿está Ana neurótica? ¿es Pedro un buen conductor? • Validez predictiva: Indica el grado en que las puntuaciones de un test predicen conductas futuras. El coeficiente de correlación obtenido entre el test y el criterio, tras un período razonable de tiempo, equivale al coeficiente de validez predictiva. Un test con validez predictiva permite contestar preguntas como ¿es posible que Ana se vuelva neurótica? ¿Es posible que Pedro sea un buen conductor?

	Validez de Constructo (teórica, estructural, conceptual) Grado en el cual una técnica mide un determinado rasgo o constructo hipotético. • <i>Validez Convergente:</i> Expresa el grado en que una medida está sustancialmente relacionada con otras medidas del mismo constructo y con otras variables con las que, por razones teóricas, debería relacionarse. • <i>Validez Divergente o Discriminante:</i> Expresa el grado en que una medición presenta una baja correlación con otras mediciones que, utilizando el mismo método, miden otros constructos. Últimos desarrollos del concepto de validez: Estándares del APA (1999) Los estándares del APA (1999) presentan una teoría unificada de la validez en la que el constructo es el eje central de la validación y en la que incorporan el énfasis en las consecuencias del uso de los test, el papel de la teoría de la generalizabilidad y el uso de los test para la toma de decisiones políticas, y definitivamente aportan la siguiente definición: "Validez es un concepto unitario. Es el grado en que todas las evidencias acumuladas apoyan la proyectada interpretación de las puntuaciones del test implicadas en el propósito propuesto... Validez se refiere al grado en que evidencia y teoría soportan las interpretaciones de las puntuaciones del test implicadas en los usos propuestos del mismo... El proceso de validación implica acumulación de evidencias que proporcionan una base científica sólida para las interpretaciones que se proponen. Esto es, deben ser evaluadas las interpretaciones de las puntuaciones del test requeridas para usos propuestos, no el test mismo". Así, el planteamiento general de los estándares de 1999, vigentes en la actualidad, no sólo reafirma la posición central de la validez, sino que entiende que los principios de la validez se basan en las inferencias sobre el significado de comportamientos o de atributos consistentes. Las principales fuentes de evidencias son (Fernández-Ballesteros y Calero, 2004): 1. <i>Evidencias basadas en el contenido del test:</i> incluye evidencias de relevancia, representatividad y calidad técnica del instrumento. Estas proceden del análisis de las relaciones entre el contenido del test y el constructo que se pretende medir. Por contenido se entiende los temas, palabras, formatos de ítems y/o tareas que forman el test tanto como las instrucciones y los procedimientos de administración y puntuación. 2. <i>Evidencias basadas en los procesos de respuesta (validez sustantiva):</i> en éstas se trata de recurrir al análisis teórico y empírico de los procesos de respuesta de las personas que cumplimentan el test para valorar el ajuste entre el constructor y la respuesta real de los examinados. 3. <i>Evidencias basadas en la estructura interna:</i> tratan de analizar las relaciones de los ítems de test entre sí y de ellos con el constructor que sirve para la interpretación de las puntuaciones obtenidas. 4. <i>Evidencias basadas en la relación con otras variables (validez externa):</i> se incluyen aquí los análisis de relaciones entre las puntuaciones del test y otras variables externas utilizadas como criterio, sean ejecuciones o puntuaciones de otros tests que midan el mismo constructo o constructos diferentes. 5. <i>Evidencia de generalizabilidad:</i> serían aquellas que examinan la extensión en que las puntuaciones y su interpretación son generalizables a través de grupos, contextos y tareas incluyendo relaciones test-criterio.

	6. <i>Evidencias sobre las consecuencias del uso del test:</i> se trata de atender a las consecuencias previstas y no previstas del uso del test, potenciales y reales, positivas y negativas.

Cuadro 20. La teoría de la generalizabilidad

LA TEORÍA DE LA GENERALIZABILIDAD
Generalizabilidad: Grado en el cual las puntuaciones halladas mediante un instrumento son generalizables a distintos ámbitos o universos de generalización.
Universos de Generalización Universo de generalización de las puntuaciones: Indica la medida en que los datos obtenidos por un observador son generalizables a los obtenidos por otro (equivale al concepto de fiabilidad interjueces de la psicometría clásica). Universo de generalización de los elementos: Indica el grado en que los ítems de un instrumento comprenden una muestra representativa del universo de conductas que se pretenden medir (equivale a validez de contenido y consistencia interna). Universo de generalización temporal: Indica el grado en que los datos obtenidos en un momento determinado son inferibles a otros momentos de la vida del sujeto (equivale al concepto fiabilidad test-retest). Universo de generalización situacional: Indica el grado en que los datos obtenidos en una situación son generalizables a otras (equivale al concepto de validez ecológica). Universo de generalización de los métodos: Indica el grado en que los datos obtenidos mediante un método de evaluación son generalizables a otros que evalúan el mismo contenido (equivale al concepto de validez convergente). Universo de generalización de las dimensiones: Indica el grado en que las estimaciones de una conducta son generalizables a otras conductas (equivale al concepto de validez de constructo).

Cuadro 21. Aplicaciones de la Teoría de Respuesta al Ítem

APLICACIONES DE LA TEORÍA DE RESPUESTA AL ÍTEM
Bancos de Ítems: Conjunto de elementos que miden un campo, rasgo o área de conocimiento almacenados en un ordenador junto con sus propiedades psicométricas y sus parámetros. Conjunto de ítems que cumplen el requisito de definir con precisión la variable habilidad que se analiza. El rasgo es el banco y la validez del contenido del banco determinará el grado de generalización de sus resultados. Construcción de Tests a Medida o Tests Adaptativos Computerizados: Son un corolario de la existencia de un banco de ítems consistiendo en la selección informatizada de aquellos ítems que mejor permiten medir la habilidad de un individuo concreto. Es un test adaptado al individuo en el que cada nuevo ítem por aplicar depende del éxito o fracaso del sujeto en los ítems anteriores. Tests Referidos al Criterio: A diferencia de los test normativos, los tests referidos al Criterio son tests cuyo interés se centra en la determinación del grado en el que los sujetos dominan ciertas materias o áreas de conocimiento, a las que se suele denominar criterio o dominio.

Equiparación de Puntuaciones: Es el proceso que permite el desarrollo de una conversión del sistema de unidades de un test al sistema de unidades de otro. En el marco de la TRI al ser las mediciones invariables con respecto al instrumento, es decir, al ser la métrica común no es necesario.

Sesgo: Un test o ítem está sesgado contra un grupo de sujetos cuando este grupo, a pesar de tener el mismo nivel que otros en la variable medida, sale sistemáticamente perjudicado al utilizar dicho test. Bajo el enfoque de la TRI un ítem estará sesgado si su Curva Característica del Ítem es distinta para los grupos evaluados. Así, estiman las CCI para ambos grupos y ven en qué medida difieren. No obstante, los métodos basados en la TRI requieren grandes tamaños muestrales para garantizar la adecuación en el funcionamiento de estos métodos.

Cuadro 22. Garantías éticas en la evaluación psicológica

GARANTÍAS ÉTICAS EN LA EVALUACIÓN PSICOLÓGICA
Principios Generales de Actuación del Psicólogo: Respeto por la persona, protección de los derechos humanos, sentido de la responsabilidad, sinceridad para con los clientes, prudencia en la aplicación de instrumentos y técnicas, competencia profesional, solidez de la fundamentación objetiva y científica de sus intervenciones profesionales. Formación del Psicólogo: Preparación y especialización en los métodos, instrumentos, técnicas y procedimientos que emplee. Conocer los límites de su competencia y las limitaciones de sus técnicas. Utilización Correcta de los Conocimientos Psicológicos: Prestación de servicios con la certeza de que no serán mal utilizados en contra de los intereses de las personas, los grupos, o las instituciones. Rigor Científico en la Actuación Profesional: No utilizará medios o procedimientos que no se hallen suficientemente contrastados dentro de los límites del conocimiento científico vigente. Uso del Material Psicológico y los Datos empleados: Confidencialidad de los datos recogidos en todo el proceso de la evaluación psicológica. Informe Psicológico: Claridad, Precisión, Rigor y Confidencialidad.
Directrices internacionales para el uso de los tests: Los 135 artículos que contienen las directrices internacionales para el uso de los tests, se dividen en dos apartados: 1) uso ético de los tests; y 2) utilización adecuada de los test. <i>1. Uso ético de los tests</i> 1.1. Actuar de forma ética y profesional 1.2. Asegurarse de que son competentes para el uso de los tests 1.3. Responsabilizarse del uso que hacen de los tests 1.4. Asegurarse de que los materiales del test están seguros 1.5. Asegurarse de que los resultados de los tests se tratan confidencialmente <i>2. Utilización adecuada de los tests</i> 2.1. Estimar la utilidad potencial de los tests en una situación evaluativa 2.2. Elegir tests técnicamente correctos y adecuados a cada situación

- 2.3. Prestar atención a los aspectos relacionados con el sesgo de los tests
- 2.4. Hacer los preparativos necesarios para la aplicación del test
- 2.5. Aplicar los tests adecuadamente
- 2.6. Puntuar y analizar los resultados de los tests con precisión
- 2.7. Interpretar los resultados adecuadamente
- 2.8. Comunicar los resultados de forma clara y precisa
- 2.9. Revisión de la adecuación del test y de su uso

Directrices para la traducción y adaptación de tests: 4 fases en el proceso de adaptación en las que se pueden producir errores que hay que tratar de minimizar:

1. *Contexto:* Equivalencia del constructo en las distintas culturas en las que se va a aplicar el test para poder comparar los resultados obtenidos.
2. *Construcción y adaptación:* Tener en cuenta las peculiaridades de la población a la que se va a aplicar; por lo tanto, es fundamental que aquellas personas que van a llevar a cabo la adaptación, además de tener conocimientos acerca del proceso de construcción de un test, conozcan los idiomas correspondientes a las distintas poblaciones y las peculiaridades culturales de cada una de ellas.
3. *Aplicación:* Se deben cuidar al máximo aspectos tales como las instrucciones que se den antes de la aplicación y las interacciones entre el aplicador y los examinados ya que pueden influir en la fiabilidad y validez del test. Los aplicadores deberán seleccionarse entre las personas pertenecientes a la población a la que se aplica el test; de esta manera estarán familiarizados con los distintos matices propios de dicha cultura; por otra parte es conveniente que reciban un cierto entrenamiento sobre la forma de aplicación, haciéndoles ver la importancia que tiene el seguir al pie de la letra las instrucciones de aplicación.
4. *Interpretación de las puntuaciones:* Para poder interpretar adecuadamente los resultados obtenidos al aplicar un test, tanto si es un test original como cualquier adaptación, es necesario tener una cierta formación psicométrica; pero, en este último caso, existe además el peligro de que al comparar los resultados obtenidos en distintos grupos o países se intente hacer algún tipo de clasificación de los sujetos. Una cosa es analizar las semejanzas y diferencias entre los grupos y otra establecer cualquier tipo de clasificación, puesto que es prácticamente imposible encontrar dos comunidades que sean equiparables completamente. Para una mejor interpretación de los resultados, el manual del test deberá incluir una información exhaustiva de todo el proceso de adaptación.

Directrices internacionales para la evaluación informatizada y test aplicados por Internet: 4 objetivos fundamentales

1. *Cuidar las cuestiones tecnológicas*

- Desarrollan consideraciones de hardware y software y centran su atención en la descripción de los requisitos del sistema, la justificación del programa y las posibilidades de actualización.
- Se dirigen a asegurar la robustez del programa y señalan la necesidad de anticipar posibles fallos y dar todo el soporte técnico a los usuarios.
- Se centran en tomar en consideración los factores humanos que puedan afectar a la presentación del material y tratan fundamentalmente del diseño, el tipo de estímulos, la distinción entre instrucciones e ítems, la facilidad del usuario para volver atrás, la familiaridad de los usuarios con las demandas del programa, la compatibilidad con diferentes sistemas informáticos...
- Tratan de considerar el ajuste de las características técnicas con diferentes causas de incapacitación en los sujetos evaluados y se centran en asegurar y explicitar quiénes pueden y no pueden ser evaluados mediante ese procedimiento.
- Se interesan en el suministro de ayuda e información práctica para los usuarios. Si las directrices del programa son claras, completas, explícitas; si proporcionan práctica inicial con el procedimiento...

2. *Garantías de calidad*

- Trata de asegurar el conocimiento y la competencia para los usuarios y los sujetos evaluados.
- Se centra en los requisitos psicométricos del procedimiento.
- Se refiere a los procedimientos que son versiones de tests clásicos de papel y lápiz, y se dirige a la demostración de la equivalencia de formatos.
- Se refiere a los análisis de fiabilidad de puntuaciones y resultados obtenidos en estas versiones.
- Trata sobre si la interpretación de resultados es apropiada señalando sus limitaciones. Si las especificaciones del programa suministran documentación y/o *feedback* adicional y señalan en que criterio, reglas o procedimientos basan la interpretación...
- Pretende asegurar la igualdad de acceso para cualquier grupo, país, étnia o comunidad lingüística.

3. *Asegurar los niveles de control apropiados*

- Sobre las condiciones de administración.
- Sobre las condiciones de supervisión.
- Sobre la práctica inicial del evaluador y la exposición inicial del evaluado.
- Sobre las posibilidades de autenticación y estafa del acceso al sistema.

4. *Asegurar la privacidad y seguridad de los datos recogidos*

- Seguridad para el material que constituye el test.
- Seguridad para la transferencia de datos por Internet.
- Confidencialidad de los resultados obtenidos.

CAPÍTULO 7.

TÉCNICAS DE RECOGIDA DE INFORMACIÓN EN EVALUACIÓN PSICOLÓGICA: ENTREVISTA, AUTOINFORMES, OBSERVACIONALES, PROYECTIVAS, SUBJETIVAS, OBJETIVAS, Y PSICOMÉTRICAS

- 7.1. Introducción
- 7.2. La entrevista psicológica
 - 7.2.1. Definición de la entrevista psicológica
 - 7.2.2. Clasificaciones de la entrevista
 - 7.2.3. Objetivos de la entrevista
 - 7.2.4. La entrevista como instrumento del proceso de evaluación-intervención psicológica
 - 7.2.5. Características y habilidades del entrevistador
 - 7.2.6. Características psicométricas de la entrevista, fuentes de distorsión y estrategias facilitadoras
- 7.3. Los autoinformes
 - 7.3.1. Definición y clases de variables evaluadas mediante autoinformes
 - 7.3.2. Principales tipos de autoinforme
 - 7.3.3. Fuentes de error de los autoinformes
- 7.4. Las técnicas observacionales
 - 7.4.1. Definición y objetivos
 - 7.4.2. Unidades de observación y de medida
 - 7.4.3. Técnicas de registro
 - 7.4.4. Consideraciones muestrales
 - 7.4.5. Fuentes de error y requisitos científicos de la observación
- 7.5. Las técnicas proyectivas
 - 7.5.1. Definición e hitos históricos de las técnicas proyectivas
 - 7.5.2. Características de las técnicas proyectivas
 - 7.5.3. El concepto de proyección
 - 7.5.4. Clasificación de las técnicas proyectivas
 - 7.5.5. Garantías psicométricas y futuro de las técnicas proyectivas
- 7.6. Las técnicas subjetivas
 - 7.6.1. Definición y características de las técnicas subjetivas
 - 7.6.2. Técnicas subjetivas representativas
- 7.7. Las técnicas objetivas
 - 7.7.1. Definición y características de las técnicas objetivas
 - 7.7.2. Tipos de técnicas objetivas
 - 7.7.3. Validez y fiabilidad de las técnicas objetivas
- 7.8. Las técnicas psicométricas
 - 7.8.1. Definición y características de los tests psicométricos
 - 7.8.2. Clasificación de los tests psicométricos
 - 7.8.3. Presente y futuro de los tests psicométricos
- 7.9. Técnicas de recogida de información en los programas de la asignatura evaluación psicológica en las facultades de psicología españolas.
- 7.10. Hacia una pluralidad técnico-metodológica en evaluación psicológica
- 7.11. Resumen. Técnicas de recogida de información en evaluación psicológica

7.1 INTRODUCCIÓN

Toda técnica es una forma específica de proceder que corresponde a una convención pragmática, limitada por los medios disponibles de operacionabilidad de los conceptos en un momento determinado. La técnica se distingue del método por cuanto no existe ninguna técnica que por sí misma nos permita llegar a un fin más allá del de la medición de un concepto. Así pues un conjunto de técnicas ordenadas en un conjunto de reglas para obtener un fin es lo que caracteriza el método. Es decir, las técnicas son las estrategias concretas y puntuales que se utilizan para operativizar el objetivo.

No obstante, con respecto a la terminología empleada para designarlas, el acuerdo no ha sido tan unánime. Así, vemos que se utilizan a veces de forma sinónima términos como tests, técnicas o instrumentos para referirnos a estos procedimientos de recopilación de información. La palabra test etimológicamente procede de la palabra latina testis, que significa prueba o ensayo (Del Rincón, Arnal, Latorre y Sans, 1995), siendo en su procedencia un concepto muy genérico. Por ello, diversos autores como Pelechano nos brindan unas matizaciones que son necesarias para clarificar estos conceptos. Instrumentos y técnicas serían vocablos equiparables que comparten la característica común y diferencial con respecto a los tests de que son procedimientos no cuantificados ni tipificados. El test, por el contrario, es un instrumento sistemático y tipificado que compara la conducta de dos o más personas (Pelechano, 1976).

Cronbach (1990) define el test como un procedimiento sistemático para observar la conducta y describirla con la ayuda de escalas numéricas o categorías establecidas...el testólogo recoge información preguntando y observando a todas las personas de la misma manera, en la misma o comparables situaciones. Esta definición engloba los cuestionarios de personalidad, procedimientos para observar la conducta social, aparatos de medida de la coordinación o, incluso, registros de productos.

Fernández-Ballesteros y Maciá (1992) definen los tests como “procedimientos de medida cuyo material, forma de administración, puntuación, corrección y valoración son estándar y de los que se derivan puntuaciones normativas”. En la misma dirección, la Asociación Internacional de Psicología define el test como “una prueba determinada, que implica una tarea a cumplir, idéntica para todos los sujetos examinados, con una técnica precisa para la apreciación del éxito o el fracaso, o para la clasificación numérica de los aciertos” (Del Rincón et al., 1995).

Recientemente, Urbina (2007) define el test como un procedimiento sistemático para obtener muestras de conducta significativas del funcionamiento cognitivo o afectivo de la persona y para puntuar y evaluar estas muestras de acuerdo con normas tipificadas. En definitiva, *los tests son procedimientos de medida cuyo material, forma de administración, puntuación, corrección y valoración son estándar y de los que se derivan puntuaciones normativas. Así, las técnicas o instrumentos se diferencian de los tests por no cumplir alguna de las características mencionadas* (Fernández-Ballesteros y Calero, 2004).

Todo intento taxonómico admite multiplicidad de variantes y la tecnología evaluativa no constituye una excepción. Al buscar un sistema clasificatorio de las técnicas nos encontramos con numerosas clasificaciones atendiendo a diferentes criterios. Así, se pueden clasificar según sus características formales, según sus características de contenido y según el ámbito concreto que exploran.

Una de las primeras taxonomías agrupaba las técnicas de medida en tres categorías: tests proyectivos, tests objetivos y cuestionarios, que actualmente resulta muy general. Posteriormente, Pervin (1979) efectuó una taxonomía, con pocas variantes, en base a cuatro categorías: tests proyectivos, tests subjetivos, tests psicométricos y tests objetivos. Si bien ofrece una clasificación que tiene un mayor poder aglutinador y ha gozado de buena acogida, sin embargo, no recoge suficientemente la diversidad de procedimientos utilizados en la actualidad, especialmente los provenientes de la evaluación conductual.

Así mismo, en la clasificación de Pervin, se emplea la palabra *test* en un sentido amplio, no el estandarizado que, en un lenguaje preciso, debiera atribuírsele. Los *tests proyectivos* serían el ejemplo más claro de pruebas no estructuradas (múltiples posibilidades de respuesta) y disfrazadas (ocultamiento del propósito del test). El test de Rorschach es uno de los más representativos dentro de esta taxonomía, juntamente con el TAT, el TRO y los tests gráficos. Suelen presentar un grado bajo de fiabilidad y validez (debido en parte a los diferentes marcos teóricos interpretativos de los evaluadores) crítica ésta de las más reiteradamente formuladas a este tipo de pruebas, juntamente con la referente al abuso inferencial.

Las *técnicas subjetivas* son no estructuradas (multiplicidad de respuesta), no disfrazadas (el sujeto conoce su propósito) y voluntarias. Según Pervin (1979) el objetivo que persiguen es obtener datos sobre las percepciones que el sujeto posee de sí mismo y de las diversas circunstancias ambientales. Es decir, el sujeto se auto-califica y califica al mundo. El autor incluye en esta categoría la entrevista, estrategia que por su importancia dentro del campo de la tecnología evaluativa es tratada como técnica independiente y con características propias por otros autores (Fernández-Ballesteros, 1983, 1992, 2004; Garaigordobil, 1998; García Marcos, 1983; Kline, 1985; Moreno 2003), las autobiografías, la técnica Q de Stephenson (1953), el diferencial semántico de Osgood et al. (1957) y el test de constructos personales (REP) de Kelly (1955). Por su parte Fernández-Ballesteros (1992) añade a las ya citadas la técnica de listado de adjetivos (LA), aunque no incluye las técnicas subjetivas como entidad diferenciada en su última clasificación (Fernández-Ballesteros, 2004).

Los *tests psicométricos* se caracterizan por ser técnicas estructuradas (número limitado de respuestas) y voluntarias (libertad en la emisión de la respuesta) con base psicométrica factorialista principalmente (aunque también se emplean estrategias racionales y empíricas). Se trabaja con material estandarizado y tipificado tanto para su aplicación, como corrección e interpretación. Desde el punto de vista histórico constituyen los primeros intentos científicos de diagnóstico, por lo que se les suele considerar los tests prototípicos, que aunque criticados han sido y son ampliamente utilizados. El cuestionario 16 Pf de Catell, el EPI de Eysenck y el MMPI de Hathaway y Mckinley, son algunos de los más representativos entre otros.

Algunos autores como Fernández-Ballesteros (1983, 1992) incluyen este tipo de instrumentos entre los autoinformes, ya que suponen un informe verbal sobre algún tipo de manifestación conductual, y son considerados como autoinformes tipificados contruidos a través de procedimientos psicométricos que permiten obtener una puntuación diferencial en una determinada variable. En este sentido, la principal diferencia con respecto a los autoinformes desde la perspectiva conductual radicaría en la interpretación de los resultados como signo de entidades psíquicas subyacentes o como muestras representativas de la conducta. Otros aspectos diferenciales se asientan en el hecho de que los autoinformes psicométricos supongan medidas indirectas de un atributo, se hayan contruido a partir de diseños intersujeto, con estrategias racionales, empíricas o factoriales y se basen en la generalidad de la conducta. En

la última revisión que realiza (Fernández-Ballesteros, 2004), los autoinformes psicométricos (MMPI, 16 PF...) se incluyen tanto en la categoría de autoinformes (Fernández-Ballesteros, 2004d) como en el capítulo de los tests psicométricos (Calero, 2007, Calero y Padilla, 2004).

Los *tests objetivos* abarcarían la categoría de pruebas estructuradas y objetivas (emisión de respuesta correcta), vinculadas al enfoque experimental. Una característica esencial es que el sujeto no puede modificar voluntariamente su respuesta, juntamente con la objetividad de puntuación. Cronbach (1960) los denomina tests de ejecución y Eysenck (1960) tests de conducta. En esta categoría se encuentran los tests que estiman las variables fisiológicas, pruebas motóricas, perceptuales, simuladores de la vida cotidiana, y tests de rendimiento-ejecución (Fernández-Ballesteros y Calero, 1992). Pervin (1979) incluyó la observación directa en situaciones artificiales y Kline (1985) considera que los tests proyectivos podrían incorporarse en esta categoría si variaran su forma de puntuación. Otros autores incluyen dentro de esta categoría los tests calificados por Cronbach (1960) de rendimiento máximo; es decir, los típicos tests de inteligencia, por cuanto se trata también de pruebas estructuradas (limitadas alternativas de respuesta) y objetivas (requieren la emisión de una única respuesta correcta). Desde el punto de vista conceptual, no obstante, hay importantes diferencias entre los tests de inteligencia tradicionales y los tests propiamente objetivos. Mientras que estos últimos se apoyan en la teoría del aprendizaje y en los presupuestos de la psicología cognitiva, adhiriéndose al método experimental, los primeros se enraízan con la teoría de rasgos y con la metodología correlacional-factorial. De aquí que algunos prefieran incluirlos entre los tests psicométricos.

Como señalaba al comienzo del capítulo, a la hora de clasificar las técnicas se pueden utilizar distintos parámetros. Fernández-Ballesteros (1992c) propuso 6 metodologías básicas en evaluación psicológica, en las cuales la observación es el método básico:

- Técnicas de observación propiamente dichas (más o menos sistemáticas y estructuradas).
- Técnicas de autoinforme que se apoyan en la introspección y autoobservación del propio sujeto a evaluar.
- Técnicas de entrevista considerada como recogida de autoinformes del propio sujeto o de personas allegadas.
- Técnicas subjetivas, bien de calificación o bien de clasificación según atributos o descripciones verbales que el sujeto realiza sobre sí mismo, o sobre su ambiente.
- Técnicas proyectivas, de carácter verbal, gráfico o constructivo.
- Técnicas objetivas cuyos datos son obtenidos a través de instrumentos y aparatos que permiten registros objetivos sin que medie la opinión del evaluador.

Sin embargo, esta autora en su última revisión (Fernández-Ballesteros, 2004), en la sección de métodos de recogida de información incluye: la observación, las técnicas objetivas, los autoinformes, la entrevista, las técnicas proyectivas, y las técnicas psicométricas o tests. Entre otras clasificaciones recientes destacar la realizada por Caballo (2005) que en lo que se refiere a las estrategias de evaluación en psicología clínica, realiza la siguiente clasificación:

- La entrevista
- Cuestionarios, inventarios, escalas
- La observación conductual

- La evaluación cognitiva
- El autorregistro
- La evaluación por los demás
- Las medidas análogas (de lápiz y papel, de audio, de video, de actuación, de representación de papeles)
- La evaluación psicofisiológica

Así mismo, se puede destacar la reciente clasificación de las técnicas de evaluación psicológica de Calero (2007) que las clasifica en:

- Técnicas objetivas - Observación
- Técnicas objetivas - Estructuradas: técnicas objetivas, registros psicofisiológicos, tests psicométricos de ejecución.
- Técnicas voluntarias estructuradas: autoobservación, autoinformes psicométricos
- Técnicas voluntarias - los autoinformes: entrevista.
- Técnicas voluntarias no-estructuradas: técnicas subjetivas
- Técnicas objetivas no-estructuradas: técnicas proyectivas

Todas estas técnicas implican procedimientos más o menos estructurados de recogida de información en Evaluación. Utilizaremos la palabra técnica con la finalidad de abarcar los múltiples procedimientos que se emplean en Evaluación Psicológica, para formular hipótesis y/o verificarlas. A pesar de que los tests son los instrumentos de recogida de información más conocidos, estos son sólo una parte pequeña de los dispositivos existentes para recoger información.

El establecer una clasificación que agrupe las distintas técnicas dentro de categorías supone, entre otros aspectos, que éstas deben ser exhaustivas y excluyentes, condiciones que son muy difíciles, si no imposibles de cumplir. Teniendo en cuenta que cualquier clasificación de las técnicas tendrá cierto nivel de dificultad por la complejidad de la tarea que supone clasificar el amplio conjunto de herramientas de evaluación existentes, en este proyecto docente, aún con las limitaciones que implica, he optado por la siguiente clasificación:

1. La entrevista psicológica
2. Los autoinformes
3. Técnicas observacionales
4. Técnicas proyectivas
5. Técnicas subjetivas
6. Técnicas objetivas
7. Técnicas psicométricas

7.2. LA ENTREVISTA PSICOLÓGICA

La entrevista, en sus distintas modalidades y tipologías, es considerada una estrategia fundamental en la Evaluación Psicológica y utilizada de hecho en todas las áreas de aplicación y desde todos los modelos teóricos existentes. Es una estrategia a utilizar no sólo durante el proceso de evaluación, también durante el proceso de intervención. La entrevista es la matriz dentro de la cual tal proceso (proceso de evaluación-intervención) es llevada a cabo. Si bien la Evaluación Conductual, en sus comienzos, desestimó la utilización de la entrevista debido a sus problemas metodológicos, a partir de los trabajos de Kanfer y Saslow (1965, 1969), que revalorizaron su importancia, se ha convertido también para el enfoque conductual en una técnica fundamental e indispensable. La expansión en el uso de la entrevista es grande, utilizándose en el campo de la evaluación e intervención clínica con niños y adultos, en la investigación social y en el campo aplicado de las organizaciones.

7.2.1. DEFINICIÓN DE LA ENTREVISTA PSICOLÓGICA

La entrevista ha sido definida de muchas maneras. Son clásicas las definiciones de Rogers y Sullivan. Rogers (1942) destacó la importancia de las emociones: “La entrevista es el artificio destinado a estimular el intercambio emocional en el que el consultor auxilia al cliente en descubrir y dominar sus propios sentimientos”, mientras que Sullivan (1954/1977) consideró que la misma estructuración de la entrevista permite elucidar las pautas del sujeto: “Entrevista es una situación de comunicación primariamente vocal, en un grupo de a dos, más o menos voluntariamente integrado sobre una base progresivamente desarrollada de experto-cliente con el propósito de elucidar pautas características del sujeto entrevistado... La entrevista es considerada como un fenómeno interpersonal, y los datos para su estudio y comprensión deben ser derivados de la observación sobre lo que ocurre entre los participantes de la entrevista o, para decirlo de otra manera, de la observación del campo de su interacción”.

Bleger (1964/1985) y Matarazzo (1965) también pusieron el énfasis en el aspecto interaccional. Para Bleger (1964/1985): “la entrevista consiste en una interacción humana en la cual uno de sus integrantes debe tratar de saber lo que está pasando en la misma y debe actuar según ese conocimiento... el entrevistador y el entrevistado constituyen un grupo, una totalidad, en el cual sus integrantes están interrelacionados y en el que la conducta de ambos es interdependiente”. Así, desde esta perspectiva, la conducta de uno es estímulo para la conducta del otro y viceversa, por lo que el análisis de la entrevista requerirá centrar el estudio sobre el entrevistador, el entrevistado, y la relación interpersonal que se establece entre ambos. Por su parte, Matarazzo (1965) definió la entrevista como un tipo de conversación, en la que dos personas, o más de dos, interactúan de forma verbal y no verbal con el propósito de cumplir un objetivo previamente definido.

Al analizar aspectos comunes en diversas definiciones planteadas de la entrevista, se puede enfatizar que la entrevista requiere: 1) Una relación entre personas (dos o más de dos); 2) Una vía de comunicación simbólica bidireccional, preferentemente oral; y 3) Unos objetivos prefijados y conocidos, así como una asignación de roles que significa, al menos idealmente, un control de la situación por parte del entrevistador. Fijación de objetivos y control son los fundamentos para que se hable, en la entrevista, de una “relación interpersonal asimétrica”.

Silva (1992) en un esfuerzo de síntesis por delimitar las ventajas y limitaciones que posee la entrevista como instrumento de evaluación subraya que entre sus ventajas se pueden destacar: (1) La relación interpersonal que implica; (2) La flexibilidad que permite el entrevistador de adaptarse sobre la marcha a las peculiaridades del entrevistado; (3) La posibilidad de observación del comportamiento del entrevistado, tanto en lo que se refiere a la información verbal que transmite, como a los aspectos no verbales; (4) La posibilidad de registrar grandes cantidades de información de tipo subjetivo (emociones y otros aspectos de la subjetividad) y de tipo biográfico (anamnesis e historia vital, situación actual de vida, expectativas futuras...); y (5) La posibilidad de evaluar a personas con trastornos mentales o escasa escolarización que no podrían ser examinadas con otros instrumentos como tests o cuestionarios. Respecto a sus limitaciones Silva (1992) enfatiza el coste relativamente elevado por el tiempo y esfuerzo que requiere por parte del entrevistador, y la interferencia de sesgos que pueden tener variada procedencia: del entrevistador, del entrevistado, de la situación, y de la relación.

La entrevista es el instrumento de evaluación que precede a cualquier modalidad de intervención o proceso de toma de decisiones, adopta un formato interactivo y está presente en el continuo evaluación-intervención (Casullo y Márquez, 2003). Es un procedimiento ampliamente utilizado con distintos objetivos o propósitos: conseguir información u ofrecer información en la evaluación e intervención psicológica; indagar sobre características de personalidad, estrategias, incidentes críticos, habilidades, competencias, y datos biográficos en procesos de selección; indagar sobre los aspectos motivacionales, intereses y competencias en los procesos de orientación; investigar el conocimiento que pone en juego un experto cuando resuelve problemas de su competencia; orientar o asesorar y comunicar a clientes o pacientes los resultados de cualquier proceso evaluativo. Es un instrumento destinado a recoger datos mediante el autoinforme o heteroinforme del sujeto evaluado y de ahí su importancia, sus fortalezas y sus debilidades (Márquez, 2004).

7.2.2. CLASIFICACIONES DE LA ENTREVISTA

En la literatura se encuentran muy distintos tipos de clasificación de las entrevistas, dependientes de los distintos criterios tenidos en cuenta para agruparlas. Entre los sistemas unidimensionales de clasificación están, los que clasifican las entrevistas según los objetivos, según el grado de estructuración de la entrevista, y según el marco teórico del evaluador.

7.2.2.1. Clasificación en función de los objetivos

La diferenciación de las entrevistas en función de los objetivos ha sido una de las clasificaciones más antiguas. Si se atiende a los objetivos, Pelechano (1976) las ha clasificado en entrevistas de investigación y entrevistas de intervención. Dentro de las entrevistas de intervención suelen diferenciarse entre entrevistas diagnósticas o de evaluación y entrevistas terapéuticas. Con el mismo eje de clasificación Avila Espada (1990a) enumera 5 tipos de entrevista: la entrevista como técnica de evaluación o diagnóstico, la entrevista de consejo u orientación, la entrevista terapéutica, la entrevista de investigación, y la entrevista focalizada. Parraga (1986) presenta una clasificación más simple distinguiendo 3 tipos de entrevista: Entrevista para obtener información, para facilitar información, y para influir en ciertos aspectos de la conducta.

7.2.2.2. Clasificación en función del grado de estructuración

En lo que se refiere al grado de estructuración de las entrevistas, eje distintivo de mayor interés en la actualidad, hay que aclarar que tal estructuración puede referirse a distintos aspectos. Como han señalado diversos autores (Fisseni, 1990) estos aspectos pueden ser: (1) preguntas, (2) respuestas, (3) realización de la entrevista incluida la fijación de una determinada secuencia de preguntas, (4) registro y elaboración de la información, e (5) interpretación de la información. Puesto que de esta manera aparecen una gran cantidad de combinaciones posibles, queda claro que no puede hablarse de estructuración o no estructuración dicotómicamente, sino de grados dentro de un continuo.

El concepto de *entrevista no estructurada* es un concepto límite, ya que es difícil que un entrevistador actúe sin esquemas, al menos implícitos, y sobre todo cuando se trata de tareas de evaluación. La expresión se utiliza más bien cuando se intenta ir al hilo de la exposición del cliente, limitándose el entrevistador a reflejar sus verbalizaciones, sin ser directivo. Este tipo de entrevista es propia de enfoques psicodinámicos y fenomenológicos.

Como veremos posteriormente al analizar la fiabilidad y validez de la entrevista se ha constatado que la entrevista está sometida a numerosas fuentes de variabilidad que hacen que la información difiera enormemente. Entre las fuentes de variabilidad, Ezpeleta (1995a) cita: "la variabilidad de la información o variabilidad en la forma de buscar y obtener la información; la variabilidad del sujeto o variación en el contenido y en el estilo de contestar; la variabilidad de criterio o variabilidad debida al uso de diferentes criterios diagnósticos; la variabilidad de la situación o desacuerdo asociado con la obtención de información diferente del mismo individuo en diferentes circunstancias; y la variabilidad de la observación o desacuerdo asociado con circunstancias en las que los observadores perciben el mismo fenómeno de forma diferente. Todas estas discrepancias son una amenaza para la fiabilidad de la entrevista". Precisamente para paliar algunas de estas fuentes de variabilidad se han creado entrevistas estructuradas y semiestructuradas que se caracterizan por presentar una lista de conductas, síntomas y acontecimientos sobre los que se debe obtener información, las pautas para realizar la entrevista y los procedimientos para registrar los datos. En función del grado de rigidez con que están presentes estos aspectos se puede hablar de entrevistas estructuradas o semiestructuradas.

Existen una amplia gama de las llamadas *entrevistas semiestructuradas*. Unos hablan de entrevista semiestructurada cuando se trabaja con preguntas abiertas y en una secuencia prefijada; otros cuando ni el texto ni las preguntas están prefijados pero se trabaja con ayuda de esquemas o pautas de entrevista que deben ser cumplimentadas. Esta segunda acepción es la más aceptada en el campo clínico.

El *grado máximo de estructuración* se alcanza si están determinados los cinco puntos o aspectos mencionados previamente. En este tipo de entrevista estructurada, el entrevistador lleva las riendas en todo momento, hace preguntas siguiendo un formulario estandarizado y demanda respuesta. Las entrevistas altamente estructuradas han tenido un gran desarrollo con la introducción de los ordenadores en la evaluación psicológica.

El desarrollo de las entrevistas estructuradas que ha llevado a la "Información Diagnóstica General" es en nuestro país un buen ejemplo de ello (Capafóns, Sosa, Alcántud y Silva, 1986). En el campo de la evaluación psicopatológica, este desarrollo ha sido en la última década impresionante, siendo un buen ejemplo el proyecto CIDI patrocinado por la OMS en torno a una

entrevista estructurada elaborada en función del diagnóstico a través del ICD-10. Estudios como los de Janca et al. (1992) evidencian “alta concordancia entre la evaluación clínica y la evaluación CIDI en problemas de ansiedad/fobias, desórdenes depresivos y en trastornos por el uso de sustancias”.

En la década de los 80 y los 90 se han desarrollado bastantes entrevistas estructuradas, especialmente orientadas al análisis de componentes recogidos en las distintas taxonomías psicopatológicas y psiquiátricas. Se trata en todos los casos, de entrevistas estructuradas que pretenden operativizar la recogida de datos para determinar la presencia o ausencia de los criterios de inclusión o exclusión definidos para cada trastorno en los sistemas de clasificación, mediante preguntas estándar que permitan valorar la existencia de experiencias subjetivas anormales y su significación clínica (Lemos, 1995). Para ello se han diseñado diversas entrevistas, algunas de las cuales se basan en las clasificaciones psicopatológicas del DSM. Un ejemplo de estas es la entrevista clínica estructurada para el DSM-III, SCID (*The Structural Clinical Interview for the DSM-III* de Spitzer, Williams y Gibbon, 1986; Spitzer, y otros 1992), y el Programa de Entrevista Diagnóstica DIS (*Diagnostic Interview Schedule*, de Robins et al., 1981).

La entrevista clínica estructurada para el DSM-III-R (SCID-I) evalúa 33 de los más frecuentes diagnósticos del Eje I en trastornos en adultos, y el SCID-II explora los desórdenes de la personalidad. Autores como Spitzer et al. (1992) han analizado la historia y descripción de ambos instrumentos. Mientras que algunos estudios concluyen que SCID-I tiene adecuada fiabilidad encontrándose altos coeficientes en relación a depresión mayor, trastornos de ansiedad y desórdenes somatoformes (Segal et al., 1993), otros estudios que analizan la validez concurrente en SCID-II (Renneberg et al., 1992) concluyen que el grado de acuerdo entre este instrumento y los resultados en el MCMI son pobres. Posteriormente, la SCID se ha adaptado al diagnóstico con el DSM-IV utilizándose en estudios que evalúan desórdenes como el estrés postraumático (Bremner et al., 1993). Estudios con SCID-PANSS (*Structured Clinical Interview for the Positive and Negative Syndrome Scale for Schizophrenia*) dirigida a la evaluación de la esquizofrenia evidencian validez de constructo, y fiabilidad interna e interevaluadores (Lindstrom, Wieselgren y Knorrning, 1994). Actualizaciones posteriores (ver Godoy y Anarte, 2004), han sistematizados guías para la entrevista clínica estructurada para los trastornos de la personalidad del Eje II del DSM-IV (First, Gibbon, Spitzer, Williams, y Benjamin, 1999), y para los trastornos del Eje I del DSM-IV (First, Spitzer, Gibbon, y Williams 1999),

Otro ejemplo de entrevistas estructuradas son los Cuestionarios para la Evaluación Clínica en Neuropsiquiatría (*Schedules for Clinical Assessment in Neuropsychiatry*, SCAN). Los Cuestionarios para la Evaluación Clínica en Neuropsiquiatría tienden a imponerse. Se componen de un conjunto de instrumentos destinados a estudiar los problemas psicopatológicos de la edad adulta. El SCAN consta de tres componentes: la 10ª edición del examen del estado presente, la lista de comprobación de grupos de ítems, y el cuestionario de la historia clínica (Lemos, 1995).

Muchos trabajos se han dirigido a desarrollar entrevistas estructuradas analizando sus coeficientes de fiabilidad y validez. Nurnberger et al. (1994) han encontrado adecuada fiabilidad test-retest de la DIGS (*Diagnostic Interview for Genetic Studies*) con los criterios del DSM-III-R en trastornos como depresión mayor, trastornos bipolares y en la esquizofrenia. Así mismo, Schramm et al. (1993) demuestran “adecuada fiabilidad test-retest y validez de la entrevista estructurada SIS-D” (*Structured Interview for Sleep Disorders*) diseñada para evaluar trastornos

del sueño de acuerdo a los criterios del DSM-III-R, y Mackinnon et al. (1993) confirman “la validez y fiabilidad de la entrevista CAMBERRA para evaluar demencia y depresión en personas de la tercera edad”.

Un estudio dirigido a diseñar entrevistas estructuradas informatizadas fue realizado por Furukawa et al. (1995) creando el sistema COALA (*Comprehensive Assessment List of Affective Disorders*) consistente en una serie de entrevistas semiestructuradas para evaluar trastornos afectivos que es capaz de proveer polidiagnósticos, ya que “se basa en los diagnósticos relacionados con 29 sistemas diagnósticos históricos y modernos a través de algoritmos informáticos”. El COALA ha sido sometido a estudios de fiabilidad interevaluador con resultados excelentes en casi todas las variables psicopatológicas y psicosociales.

Los esfuerzos por estructurar entrevistas en diversos sentidos están entre los más notables de los últimos 20 años, habiendo logrado una importancia creciente tanto en la evaluación de adultos como de niños, y en la investigación (Bidaut-Russell et al., 1995; Edelbrock y Costello, 1990; Ezpeleta, 1990a, 1995ab; Munoz, 2003; Vizcarro et al., 1986; Wiens, 1990).

En relación a las entrevistas estructuradas para niños algunos autores (Hodges, 1993) consideran que estas entrevistas permiten obtener valiosa información sobre el estatus mental infantil. La revisión realizada por este autor indica “que los niños pueden informar fiablemente y sus informaciones pueden concurrir con el conocimiento que los adultos tienen de ellos”. Sin embargo, Hodges (1993, 1994) establece diferenciaciones entre unas entrevistas y otras. Constata validez y fiabilidad en instrumentos como CAS (Child Assessment Schedule), DICA (Diagnostic Interview for Children and Adolescents), y K-SADS (Kiddie Schedule for Affective Disorders and Schizophrenia), pero destaca falta de evidencia de fiabilidad de DISC (Diagnostic Interview Schedule for Children).

Sin embargo, a veces los resultados son discrepantes ya que Jensen et al. (1995) encuentran fiabilidad test-retest de la DISC-PC, apoyando “su utilidad en el ámbito de la clínica y en la investigación epidemiológica” y en la misma dirección Bidaut-Russell y otros (1995) encuentran que el nivel de acuerdo entre padres y adolescentes en la entrevista PC-DISC fue sustancial, pese a que también observaron un ligero grado de discrepancia en la información recogida con este instrumento desde dos fuentes de información. Así mismo, otros investigadores (Boyle et al., 1993) también hallan resultados contradictorios ya que no obtienen coeficientes aceptables de fiabilidad de DICA-R, poniendo en duda la fiabilidad de esta entrevista.

En el contexto de esta ambigüedad, existen autores (Peñate, 1993) que opinan que dentro de la evaluación clínica, uno a uno, paciente-terapeuta, “este tipo de entrevista se revela como un procedimiento poco flexible, indicándose su uso preferentemente en estudios de clasificación, epidemiológicos, aunque pueden ser de utilidad en las admisiones hospitalarias”.

Relacionado con este punto, está la introducción del ordenador como entrevistador. Si bien esta es una cuestión a debate en la actualidad, y que abordaremos posteriormente (capítulo 8), existen datos empíricos que cuestionan la equiparación de datos obtenidos de una u otra manera. No obstante, el interés que la introducción del ordenador conlleva, al menos en el campo de entrevistas “hechas a medida”, y las posibilidades de análisis de datos que ello implica, parece, sin embargo, incuestionable.

7.2.2.3. Clasificación en función del marco teórico del entrevistador

Finalmente, y como un aspecto estrechamente vinculado al grado de estructuración de la entrevista, se podrían diferenciar las entrevistas clínicas en función del marco teórico que subyace al entrevistador, ya que éste condicionará las características que adopte la entrevista.

La entrevista fenomenológica focaliza su atención en las características y actitudes que debe poseer el entrevistador (autenticidad, aceptación incondicional del cliente, comprensión empática y capacidad de transmitir esta comprensión) potenciadoras de la creación de un clima terapéutico que, centrado en el cliente y en su capacidad de reorganizar su vida, constituyan el motor del cambio. Desde la perspectiva fenomenológica se comparte la creencia de que el proceso de crecimiento se da, si el entrevistador o terapeuta es capaz de establecer una relación de ayuda en un clima terapéutico adecuado, que facilite la liberación del impulso hacia la madurez, hacia la autoactualización. Se emplea un modelo de entrevista no directiva y libre, en la que el entrevistado marca las pautas del desarrollo de la entrevista, mientras que el entrevistador trata de reflejar lo manifestado por el entrevistado, y conduce no directivamente al entrevistado al análisis de sus percepciones y sentimientos, al proceso de actualización y de cambio.

La entrevista psicodinámica, a la que subyacen las teorías psicodinámicas, se centra en la elaboración personal de las experiencias tempranas como determinantes de la situación actual, ya que se comparte la creencia de que las experiencias infantiles tienen gran importancia en la construcción psíquica o de la personalidad e influyen en la vida adulta. Desde esta perspectiva se emplea un modelo de entrevista libre y semiestructurada, usando como técnica la asociación libre y el uso de interpretaciones.

La entrevista conductual vinculada al modelo conductual, focaliza su atención en la identificación de problemas concretos que presenta el entrevistado y se esfuerza por operacionalizarlos a nivel de conductas específicas, indagando sus características topográficas (frecuencia, intensidad, duración,...), así como sus relaciones funcionales con características ambientales externas o internas (antecedentes y consecuentes). Al partir del postulado de que los problemas conductuales del sujeto son aprendidos, también se analizan las características situacionales (externas e internas) que estaban presentes en el momento de la adquisición de las conductas-problema, así como los cambios ocurridos en función de cambios ambientales. El modelo de entrevista empleado es directivo y altamente estructurado.

7.2.3. OBJETIVOS DE LA ENTREVISTA

7.2.3.1. Objetivos de la entrevista inicial

No hay ninguna técnica a la que se le hayan asignado tantos objetivos como a la entrevista. Los objetivos de la entrevista inicial varían de algún modo en función del modelo teórico del profesional, pero existe un acuerdo en considerar que deberían tenerse en cuenta los siguientes (Maganto y Cruz, 2003):

- Guiar la entrevista al objetivo establecido.
- Establecer un buen rapport, es decir, un clima de confianza apropiado para propiciar la comunicación del paciente.

- Percibir al paciente tal y como es, atendiendo a su conductas verbal y no verbal y registran el grado de coherencia entre ambas.
- Contener la angustia y ansiedad del entrevistado.
- Escuchar, es decir, dedicar más tiempo a atender a lo que paciente dice, sin interrumpir, sin juzgar, etc... que a preguntar. Esto conlleva interesarse por lo que cuenta, aunque esto pueda parecer obvio, es decir, por el contenido de lo que expresa, con las características del lenguaje, el modo de informar, los silencios...
- Obtener información, lo más exacta y válida posible, sobre el problema o problemas que presenta el paciente.
- Estimular la expresión verbal del paciente, realizando preguntas adecuadas para obtener datos necesarios
- Captar lo que el paciente está solicitando a través de sus problemas, y qué suscita en el entrevistador.
- Lograr una comprensión del problema, es decir, hacer insight con el problema del consultante. Esto va más allá de la mera observación y registro intelectual de datos. Como se explicará más adelante, esto supone integrar los datos aportados, darles, sentido, conocer los antecedentes y consecuentes del mismo, etc.
- Conocer las tentativas de solución que se han implementado hasta el momento y los resultados obtenidos de las mismas.
- Establecer hipótesis diagnósticas en los términos que cada profesional entienda en función de su marco teórico. En general, las Clasificaciones nosológicas más comunes son la referencia diagnóstica para el psicólogo, pero puede operarse con otras alternativas teóricas que sean punto de referencia teórica para el profesional.
- Planificar el proceso psicodiagnóstico. Antes de finalizar la entrevista el psicólogo debe comunicar las condiciones de trabajo a seguir: horario, duración de las sesiones, técnicas complementarias, objetivos a alcanzar, exámenes complementarios de otros profesionales, cuestiones económicas si fuera el caso, etc.
- Organizar un mapa conceptual sobre los problemas que le han consultado. Tras finalizar la entrevista, conviene que el profesional dedique un tiempo a reflexionar y elaborar una representación gráfica o mapa conceptual que estructure el caso y aporte una explicación y comprensión del mismo.

Indudablemente una de las tareas primordiales del psicólogo en la primera entrevista es conocer y comprender el problema del paciente. Las estrategias que utilice para llevar a cabo esta investigación dependen, como se ha dicho en otros apartados, no sólo del modelo teórico con el que trabaja, sino de sus conocimientos, habilidades técnicas y experiencia profesional, sin desestimar al propio paciente. Para Busse y Rybsky-Beaver (2000), Fernández Liria y Rodríguez Vega (2001), Giordano (1997), y Martorell y González (1997) entre otros, se sigue un proceso para definir cuál es realmente el problema del paciente y definir un diagnóstico se deben tomar en consideración los siguientes aspectos, aunque no siempre pueden concluirse en una primera entrevista. Sintetizando los objetivos de la entrevista inicial, especialmente enfocada a ámbitos clínicos, Maganto y Cruz (2003) identifican 10 objetivos (Cuadro 23).

Cuadro 23. Objetivos de la entrevista inicial (Maganto y Cruz, 2003)

OBJETIVOS DE LA ENTREVISTA INICIAL	
1.	Motivo de consulta: Identificar el problema o los problemas en todas sus dimensiones: su naturaleza.
2.	Percepción del sujeto. Conocer cómo percibe el sujeto dicho problema y las emociones vinculadas al mismo.
3.	Análisis de la demanda: Por qué acude ahora, consultas previas, expectativas de solución.
4.	Antecedentes de su/s problema/s en su momento de aparición y en el momento actual.
5.	Consecuencias del problema Observar el grado de compromiso clínico que conllevan estos problemas para el paciente.
6.	Determinar la severidad del mismo.
7.	Tentativas de solución que el paciente ha implementado hasta el momento y los resultados obtenidos de las mismas.
8.	Jerarquizar los problemas presentados dependiendo de la gravedad de los mismos, de la urgencia en solventarlos o bien de la viabilidad del cambio.
9.	Definición operativa del problema o problemas del paciente.
10.	Hipótesis diagnósticas en los términos que cada profesional entienda en función de su marco teórico Prever las sesiones de exploración clínica que serán necesarias para terminar de formular un diagnóstico.
11.	Pronóstico. Determinar el mismo atendiendo a las variables implicadas: personales, familiares, sociales...

7.2.3.2. Objetivos de la entrevista de devolución

El objetivo de la entrevista de devolución es informar oralmente de los resultados de la evaluación psicológica. Para ello, tras el proceso de recogida de información, esta se procesa, se elabora un informe psicológico que contenga los aspectos relevantes de la evaluación (datos de identificación de evaluado; motivo de consulta y otros problemas que aparecen; áreas de funcionamiento positivo; exploración realizada (variables medidas y técnicas empleadas); comportamiento del sujeto en el proceso de la evaluación; resultados obtenidos con los instrumentos aplicados; diagnóstico (conclusiones); orientación-recomendaciones).

7.2.4. LA ENTREVISTA COMO INSTRUMENTO DEL PROCESO DE EVALUACIÓN-INTERVENCIÓN PSICOLÓGICA

En lo que a la entrevista se refiere, pueden identificarse dos momentos en los que su frecuencia es prevalente: el inicio de la evaluación y la comunicación de los resultados. En el análisis de la entrevista planteado por Márquez (2004ab) establece el siguiente esquema:

El inicio de la evaluación: La primera entrevista tendrá diferentes objetivos en función del tipo de demanda (clínica, organizacional...). Tres tipos de demanda: 1) Cuando la evaluación está relacionada con las quejas de un cliente o paciente; 2) La evaluación se realiza en contextos específicos de la práctica clínica y forense; y 3) En contextos organizativos la entrevista tiene distintas utilidades como herramienta para el análisis de puestos de trabajo, para detección de perfiles de ejecución y para diseñar programas de formación. *Los contenidos de la primera entrevista* deben someterse a un doble ajuste, por un lado, el tipo de demanda y por otro las características del sujeto. En el primer tipo de demanda, cuando se centra en una queja presentada por el sujeto o sus allegados, el contenido suele estar relacionado con

comportamientos, situaciones, percepciones, sentimientos o inquietudes que le conciernen y que implican a otros y a alguno o varios contextos en los que se desenvuelve su vida. Con este supuesto en la entrevista se suele recoger información de los siguientes aspectos: 1) del pasado (nacimiento y desarrollo, familia y entorno social del nacimiento y desarrollo, contextos educativos, experiencias laborales, condiciones biológicas en el pasado, acontecimientos vitales destacables...); y actuales (condiciones físicas del lugar donde vive, condiciones familiares, relaciones sociales, condiciones económicas, ocupación, estilo de vida, salud, intereses, motivaciones...). En el segundo tipo de demanda, realizada en la red de salud mental o en contextos de práctica forense, el estado del sujeto de evaluación es determinante para iniciar una primera entrevista o para optar por otros procedimientos o recursos (Muñoz, 2003). En este caso se trata de valorar los siguientes contenidos: atención, memoria, orientación, pensamiento, lenguaje y motricidad.... En el tercer tipo de demanda, la entrevista inicial con propósito de selección, los contenidos giran en torno a: experiencia (currículum académico y datos biográficos formativos y profesionales); características personales (aptitudes, estilos, habilidades, valores e intereses), características laborales (conocimiento y destrezas). En relación a la *realización de la primera entrevista*, tras la selección de los contenidos, estos deben ser organizados para la recogida de datos de los mismos en un formato que puede tener distintos grados de estructuración: formato estructurado, semiestructurado y no estructurado. El análisis de la información obtenida en la entrevista, que permita formular hipótesis diagnósticas

La entrevista en la comprobación de hipótesis: 1) Comprobación de hipótesis basadas en el supuesto de cuantificación; y 2) Comprobación de las hipótesis basadas en el supuesto de semejanza (exploración de las características relacionadas con los trastornos), y la comprobación del ajuste al criterio (contenidos: puntos débiles, la capacidad de interacción social, el esfuerzo adicional, la observación general del comportamiento).

La entrevista en la comunicación de los resultados: Al concluir el proceso de la evaluación es preceptiva la comunicación de los resultados a la persona o personas que la han solicitado. Los resultados deben estar recogidos en el informe. El informe se suele presentar al cliente en situación interactiva, en la se debe recordar el motivo de consulta, la evaluación realizada y los resultados obtenidos.

7.2.5 CARACTERÍSTICAS Y HABILIDADES DEL ENTREVISTADOR

La revisión realizada por Maganto y Cruz (2003), en relación a las características fundamentales del entrevistador, les lleva a considerar:

Actitudes fundamentales:

- Empatía: capacidad del entrevistador de comprender al paciente, y ser capaz de transmitir al paciente dicha comprensión.
- Calidez: actitud de cercanía.
- Competencia: el entrevistado debe recibir mensajes que le reaseguren haber consultado con un experto.
- Flexibilidad y tolerancia.
- Honestidad y ética profesional.

Habilidades de escucha

- Dejar hablar.
- Escucha activa.
- Baja reactividad verbal: el tiempo que tarda en contestar el entrevistador desde que el entrevistado ha intervenido. Una latencia prolongada favorece la expresión verbal del entrevistado.
- Silencios instrumentales: silencios que favorecen la relación interaccional y promueven que el entrevistado siga hablando.

Habilidades comunicacionales: estrategias en el manejo de las verbalizaciones

a) Estrategias para elicitación o mantener una comunicación con el paciente

- La técnica especular o de eco: tiene componentes no verbales, pero en lo verbal se expresa con una frase similar a la dicha por el entrevistado. Ayuda al entrevistado a centrarse en el tema que está tratando. Es una técnica no-inducida, ya que las verbalizaciones del entrevistador no comprometen.
- Darle la palabra.
- Comentarios confirmatorios: el entrevistador expresa aprobación, lo que alienta al paciente a continuar su discurso.
- Realimentación comunicacional: a) realimentación informativa de los hechos, se trata de repetir lo que el paciente ha dicho para asegurarnos que hemos entendido bien; y b) realimentación del comportamiento, cuando le decimos al otro las reacciones que producen sus palabras o su comportamiento.
- Señalamiento: el entrevistador pretende evidenciar un problema del paciente que él mismo ha verbalizado sin tomar conciencia de ello.
- La interpretación: es una técnica cuyo objetivo es establecer causas y consecuencias de los hechos narrados.
- Aterrizaje en paracaídas: el examinador toma conciencia de que el paciente no está planteando un tema que es necesario tratarlo, y en un momento determinado lo hace él de forma sorpresiva y directa de tal forma que el paciente tenga que encararlo inevitablemente.

b) Estrategias en el modo de hacer preguntas

- Preguntas abiertas: son aconsejables para explorar el campo del paciente.
- Preguntas cerradas: suelen ser contestadas con un monosílabo y se formulan para confirmar una información.
- Devolver la pregunta: al entrevistado formulada de otro modo, elicitación que éste siga hablando.
- Preguntas facilitadoras: no crean ambigüedad y facilitan una respuesta en una dirección.
- Preguntas clarificadoras o técnica de sondeo: se trata de preguntar con gestos o palabras cómo entiende el paciente aquello de lo que está hablando; es una forma neutra de demostrar interés y no prejuzga.

- Preguntas con encabezamiento.
- Preguntas guiadas o de respuesta inducida: la pregunta lleva implícita una respuesta monosilábica.
- Preguntas de confrontación: además de ser inductivas y guiadas confrontan al paciente con el problema que se está planteando.
- Las técnicas de presión: a) Técnica de la confrontación directa, se trata de hacer tomar conciencia al entrevistado de las contradicciones entre lo que está diciendo y su conducta no verbal, o bien entre lo que ha dicho en un momento de la entrevista y lo que dicen en otro; b) Recordar límites, presión del tiempo, se trata de observar cómo trabaja un individuo bajo la presión del tiempo, cómo organiza la información restante y se encara con el límite temporal; y c) Centrar el problema/revisión de síntomas, se trata de hacer preguntas que incidan en lo que interesa conocer.

La reflexión sobre las habilidades del entrevistador realizaba por Gongora (2005) le conduce a enfatizar las siguientes habilidades para entrevistar:

- Habilidades empáticas.
- Habilidades no verbales: voz modulada, suave y firme; animación de la expresión facial; mirar con seguridad y directamente a los ojos del paciente; asentir de vez en cuando con la cabeza; sonrisas intermitentes; gestos ocasionales con las manos; velocidad moderada del habla; empleo intermitente de la expresión “uhuh, uhuh”; cuerpo relajado; postura corporal adecuada y dirigida hacia el paciente.
- Conocerse a uno mismo.
- Habilidades para preguntar: preguntas abiertas y cerradas, peticiones de aclaración o ampliación.
- Operacionalizar las informaciones: descripciones precisas del problema.
- Parafrasear: repetir palabras o frases dichas por el paciente.
- Reflejar sentimientos: realizar descripciones de los sentimientos del paciente.
- Dirigir la entrevista: para que entrevistador y entrevistado mantengan cada uno su papel.
- Mantener la interacción: mantenimiento de la calidad y el flujo de la interacción entre terapeuta y paciente. Para ello el terapeuta debe estar atento a los comportamientos verbales y no verbales del paciente.
- Resumir: presentar al paciente una síntesis de las principales informaciones que este ha ofrecido.

Entre las actitudes relacionadas con las buenas habilidades para la entrevista Kaplan y Saccuzzo (2006) incluyen: cordialidad, autenticidad, aceptación, comprensión, franqueza, honestidad e imparcialidad.

7.2.6. CARACTERÍSTICAS PSICOMÉTRICAS DE LA ENTREVISTA, FUENTES DE DISTORSIÓN Y ESTRATEGIAS FACILITADORAS

La calidad de un instrumento se juzga esencialmente teniendo a la vista sus objetivos, y la entrevista es un instrumento al que se le han asignado objetivos muy variados (evaluación, tratamiento, investigación...). La multiplicidad de funciones implícitas en la entrevista psicológica junto a la complejidad y número ilimitado de variables que inciden en dicha situación (del entrevistador, del entrevistado, de la temática abordada, de la situación o contexto físico en el que se realiza...) han sido, sin duda alguna, los factores que han frenado el acercamiento científico que se ha producido en relación a la misma. Con objetivos tan diversos y heterogéneos no resulta fácil hacer una valoración general de los resultados sobre fiabilidad y validez de la entrevista.

En cuanto a la validez, ésta se refiere al grado en que el informe verbal del cliente durante la entrevista refleja adecuadamente su conducta en el ambiente natural. En el estudio de la bondad psicométrica de la entrevista, se ha estudiado la validez de contenido y la validez criterial.

Algunas investigaciones se han centrado en el análisis de la *validez de contenido*, en particular, en relación con el carácter más o menos completo de la información recogida a través de formatos semiestructurados o estructurados de entrevista. Los cometidos y modelos (escuelas) en los que se sustenta una entrevista influirán en los contenidos a evaluar, siendo estos contenidos más claros en el caso de entrevistas conductuales o de entrevistas diagnósticas estructuradas, y menos claros en las entrevistas de enfoque fenomenológico y no estructuradas. De cualquier forma, es un prerrequisito que se explicita en cada entrevista qué tipos de contenidos se van a evaluar y mediante qué preguntas o mediante qué estrategias. El desarrollo reciente de las entrevistas estructuradas ha mejorado este aspecto, aunque tampoco han estado libres de críticas, ya que no siempre evalúan por igual, las distintas áreas que se recogen en la clasificación psicopatológica en la que se sustenta (el caso de la DIS ejemplifica esto último con las distintas revisiones que ha tenido su contenido) (Peñate, 1993).

La *validez criterial* ha sido el criterio de validez por excelencia. En algunos casos se ha empleado el criterio de acuerdo intercalificadores como validez criterial, sin embargo, este procedimiento ha sido contestado, enfatizándose que se trata más de un procedimiento de fiabilidad que de validez. Además, se cuestiona que el acuerdo suponga presencia real de la sintomatología, ya que pudieran errar los dos evaluadores. Otra alternativa para evaluar la validez criterial ha consistido en evaluar en qué medida una entrevista puede reflejar la patología de alguien previamente diagnosticado. Los escasos datos informan que es más probable que una entrevista acierte cuando no existe patología que cuando existe, y que el acuerdo de existencia de patología es más elevado en el caso de patologías menos graves que en las más graves (Robins et al. 1981; Widiger y Frances, 1987). La heterogeneidad que reina en la información referente a la validez criterial de datos de entrevista es enorme, por lo que es difícil encontrar alguna tendencia general. Los coeficientes mejoran a medida que la entrevista se acerca a criterios más objetivos, así como a sistemas diagnósticos respaldados a su vez por mayor validez y mejor estructurados. No obstante, se constatan distorsiones según el tipo de información requerida en la entrevista (más o menos comprometida, íntima), así como según el tipo de finalidad de la entrevista (en una entrevista en el que el candidato se juega un puesto de trabajo se distorsiona más que en una entrevista clínica).

La revisión de Avila Espada (1990a) le conduce a concluir que los distintos tipos de validez de la entrevista dependen estrechamente de los objetivos de la misma, del contexto en que se realiza y el sistema teórico o marco referencial del entrevistador. Así mismo, los metaanálisis realizados por McDaniel et al. (1994) ponen de manifiesto que la validez en la entrevista depende del contenido de la entrevista, de su grado de estructuración y de la naturaleza del criterio. Además, estos autores encuentran altos niveles de validez de las entrevistas estructuradas.

En cuanto a la fiabilidad, se ha estudiado preferentemente la fiabilidad inter e intrajueces y la fiabilidad test-retest. Respecto a la fiabilidad inter e intrajueces, es decir, el grado de acuerdo entre las estimaciones y valoraciones efectuadas por distintos observadores para una misma entrevista se produce una evolución de los primeros estudios con estrategias bastante primitivas donde se daban bajos coeficientes, a posteriores estudios realizados con estrategias más sofisticadas donde los coeficientes alcanzan límites muy satisfactorios. Se encuentra que el acuerdo entre evaluadores es tanto más alto cuanto menos específico es el diagnóstico; y una clara tendencia a constatar coeficientes más altos cuanto se valoran aspectos conductuales, objetivos, y accesibles a la observación, y tanto más bajos cuanto se valoren impresiones subjetivas e interpretaciones personales (Mateu y Fuentes, 1992). La investigación ha mostrado que el mayor coeficiente de acuerdo se obtiene en relación directa con el grado de entrenamiento y experiencia de los entrevistadores en la utilización de sistemas de categorías bien delimitadas y definidas (Avila Espada, 1990a). No obstante, el acuerdo intercalificadores no debe ser tomado como un criterio de bondad psicométrica más, sino como un prerrequisito indispensable para que una entrevista pueda llevarse a cabo con garantías que superen la casuística o la fortuna en el diagnóstico. A partir del desarrollo de las entrevistas estructuradas inspiradas en algún sistema clasificatorio de los desórdenes mentales se han incrementado tanto los estudios de fiabilidad intercalificadores como los coeficientes o porcentajes de acuerdo (Peñate, 1993).

En cuanto a la fiabilidad test-retest (concordancia entre entrevistas realizadas en momentos diferentes a un mismo sujeto) la investigación ha mostrado una mayor concordancia en las apreciaciones negativas que en las positivas (Silva, 1992). Por lo que respecta a la estabilidad test-retest, algunas entrevistas estructuradas han aportado índices de concordancia en el tiempo (*Personality Assessment Schedule*, la SAP, la SADS, la DIS). En todos los casos aparecen categorías diagnósticas que obtienen índices de estabilidad aceptables, y otras que obtienen índices bastante cuestionables. El muestreo insuficiente de contenidos, el tamaño de la muestra o la propia naturaleza cambiante de lo evaluado se han esgrimido como argumentos explicativos de los bajos niveles de test-retest alcanzados.

Un grupo de estudios se ha dirigido a analizar las características psicométricas de la entrevista. Pero en cualquier caso, de la misma forma que ocurre, por ejemplo, con los autoinformes o las escalas de calificación, los criterios de bondad de un instrumento no son extrapolables a todos los instrumentos que pertenecen a la misma categoría, sino que cada uno debe someterse a la validez de su uso.

Mientras que algunos autores (García Marcos, 1983) distinguen tres aspectos que inciden directamente en los distintos tipos de validez de la entrevista: los que se refieren al entrevistador, los que se refieren al entrevistado, y los que se refieren al método, otros (Avila Espada, 1990a) clasifican los fenómenos y efectos de sesgo y distorsión en la entrevista en tres apartados: (1) los fenómenos y efectos del entrevistador; (2) los fenómenos y efectos del

entrevistado; y (3) los fenómenos y efectos de la situación de la entrevista. Mateu y Fuentes (1992) destacan factores que afectan a los sesgos que pueden producirse en una entrevista como: la habilidad del entrevistador, los estatus de ambos, los efectos de reactividad, los contenidos sobre los que se van a tratar, y el desconocimiento o la falta de información.

De los estudios sobre fiabilidad y validez de los datos de entrevista, y de la experiencia de los entrevistadores, se acumula en la bibliografía una impresionante cantidad de observaciones acerca de los peligros a evitar y de las estrategias a seguir para lograr una información de mayor calidad. Distintos autores (Bingham y Moore, 1973; Fear, 1979; Fisseni, 1990; García Marcos, 1983; Mateu y Fuentes, 1992; Parraga, 1986) detallan sesgos y sugerencias de acción para una realización adecuada de la entrevista.

García Marcos (1983) analizando varios trabajos subraya algunos errores a evitar en una entrevista conductual, tales como: Reforzar al entrevistado de forma insuficiente, utilizar los refuerzos de forma indiscriminada, emplear un alto porcentaje de preguntas cerradas, permanecer pasivo permitiendo que el entrevistado dirija la entrevista, ejercer un excesivo control y directividad durante la entrevista, sobreestimar la fragilidad del entrevistado evitando abordar áreas delicadas, intentar preguntar varias cosas a la vez, pasar por alto la comunicación no verbal que proviene del entrevistado, o interrumpir al entrevistado.

Fisseni (1990) ofrece algunas indicaciones para ser un buen entrevistador, tales como: tener un plan definido, conocer las variables de criterio, organizar la entrevista adecuadamente en relación al tiempo disponible; facilitar la motivación del entrevistado, garantizar la privacidad de la información, dejar hablar al entrevistado, ajustar el lenguaje al lenguaje del entrevistado, tener conciencia de sus propios prejuicios para evitar su influencia, registrar adecuadamente los datos obtenidos, etc.

Todas las recomendaciones que se han planteado, en general van encaminadas a mejorar la calidad de la entrevista y pueden ser vistas como criterios de calidad de la misma, es decir, de fiabilidad, validez y objetividad. Silva (1992) ordena las características de la entrevista que han demostrado mejorar la calidad de la información recogida, planteando las siguientes conclusiones:

a) Asegura una mayor objetividad:

- La estructuración de preguntas y respuestas.
- El uso de categorías de registro claramente definidas.
- La ordenación de las respuestas en esquemas o pautas.
- La unidad de perspectivas de juicio entre los distintos entrevistadores.
- El entrenamiento en el registro y la tabulación de la información.

b) Asegura una mayor fiabilidad:

- La estimación de las bases motivacionales de la respuesta.
- La utilización de racimos de preguntas, más que de preguntas aisladas.
- La clara especificación temporal de las preguntas.
- Preguntar por acontecimientos objetivos-indicadores observables de la conducta.
- Facilitar al entrevistado la posibilidad de contrastar la adecuación de sus respuestas.

c) Asegura una mayor validez:

- Dar instrucciones y aclaraciones que sitúen al entrevistado en los diversos temas.
- La transparencia en las preguntas y la utilización de un lenguaje comprensible.
- La facilitación de diversas alternativas de respuesta.
- El uso de preguntas que se circunscriban a intervalos temporales o a acontecimientos determinados.
- La preferencia de indicadores observables de la conducta.
- Los intentos de detectar y neutralizar tendencias de respuesta.
- Asegurar la confidencialidad de la información y el secreto profesional.
- Abrir las posibilidades de contrastación de las respuestas a través de otras fuentes de información.
- Evitar sugerencias.

Así mismo, Silva (1992), recomienda el uso de una serie de estrategias o técnicas especiales que ayudan a mejorar la calidad de la información recogida en la entrevista tales como la descripción de un día concreto o de acontecimientos concretos, utilización de hojas cronológicas, etc.

Respecto al proceso de entrevistar, se han planteado dos tipos de acercamiento, uno clínico-intuitivo y otro experimental-objetivo (Peñate, 1993). El primero da cuenta de una serie de guías de cómo llevar a cabo una entrevista, basadas normalmente en la experiencia clínica de los que lo proponen. Desde esta perspectiva se realizan diversas recomendaciones tales como: hacer hablar al entrevistado, saber escuchar, dirigir la conversación hacia los tópicos críticos y/o esenciales según los objetivos prefijados, e interpretar lo que el entrevistado está respondiendo según criterios situacionales en el sentido amplio. El otro enfoque, el experimental-objetivo, se dirige hacia aquellos hallazgos científicos aplicados o aplicables a la entrevista. Dentro de este acercamiento se han abordado las siguientes áreas de investigación que pueden incidir en el proceso de obtención de la información:

- a) El establecimiento de la relación terapéutica: Para que la información de la entrevista sea fiable y válida es importante que se establezca una buena relación terapéutica entre entrevistador y entrevistado. El establecimiento de unas relaciones de confianza ha sido subrayado por casi todas las corrientes y escuelas psicológicas, incluso desde el enfoque conductual-cognitivo (Prieto, 1995). Los terapeutas de conducta las consideran imprescindibles, llegando algunos autores a indicar que de todas las escuelas terapéuticas es la conductual la que considera la alianza con el cliente como un aspecto central del proceso. Lo complicado ha estado en identificar qué variables están implicadas (favorecen) esas relaciones de comprensión y confianza. Variables como honestidad, autenticidad, calidez y empatía se han revelado como facilitadoras de esa relación terapéutica.
- b) El condicionamiento semántico: La doble vía de condicionamiento clásico y operante hace que sea posible que el proceso de entrevista se pueda ver alterado en sus contenidos debido a estos elementos. Es especialmente relevante el condicionamiento operante en la entrevista en lo que se ha denominado *Efecto Greenspoon*. De acuerdo con este, las producciones lingüísticas de un individuo son susceptibles de modificarse en su tasa de ocurrencia a través de la aprobación verbal de un interlocutor. A la vista de los resultados de diversos estudios

que han demostrado que cualquier tipo de material verbal podía ser objeto de condicionamiento, el proceso de entrevista consistente no sólo en preguntas y respuestas, sino en comentarios aclaratorios o aprobatorios, atención (mirada), o gestos de comprensión, puede verse alterado por el sometimiento al condicionamiento operante de algunos de los contenidos referidos por el cliente. Al mismo tiempo el entrevistador dispone de una herramienta sencilla, la aprobación verbal o gestual, para dirigir la entrevista hacia aquellos contenidos que se consideren relevantes. Aunque hay que tener en cuenta que los refuerzos sociales también están sujetos a procesos de saciación, de forma similar a los refuerzos físicos o apetitivos.

- c) La comunicación no verbal: La comunicación en la entrevista es esencialmente verbal, aunque no exclusivamente. Las posturas corporales, los gestos o las indicaciones faciales son otra fuente importante de información que un experto entrevistador no puede dejar pasar. La información no verbal contempla no sólo los gestos o posturas (dilatación pupilar, mirada, sonrisa, movimientos de cabeza, de pies, de piernas, cambios de postura, temblores...) sino la distancia (cercanía-lejanía) y los elementos paralingüísticos (latencia, longitud de la verbalización, frecuencia de verbalizaciones, errores lingüísticos, vacilaciones, tartamudeos, entonación inadecuada...). En relación a este tema existen excelentes monografías como la de Mehrabian (1972), Argyle (1988), Peace (1981/1988) o Davis (1989). En general, los estudios se refieren a la comunicación no verbal por parte del entrevistado, sin embargo, poco se sabe sobre este tipo de comunicación en el caso del entrevistador, que no sean consejos sensatos de contacto ocular, distensión, no mostrar impaciencia, discreto contacto físico o talante amable y cortés. Evidentemente, todos estamos sujetos a las indicaciones no verbales, a su interpretación y a responder ante ellas (a veces de forma refleja), por lo que los gestos y posturas del entrevistador deben ser objeto de un control riguroso.

Recientemente, Kaplan y Saccuzzo (2006), realizan una serie de recomendaciones para manejar entrevistas interétnicas, interculturales y entre clases: 1) incrementar la conciencia cultural (ser sensible a las diferencias culturales, de clase y étnicas); 2) conocerse a uno mismo (estereotipos, prejuicios...), 3) ser flexible (evitar ideas preconcebidas, estar abierto a aceptar una perspectiva diferente a la suya), y 4) ver más allá de uno mismo (ser capaz de apreciar la perspectiva del otro). Y así mismo destacan como respuestas a evitar en una entrevista: los juicios de valor, las afirmaciones de sondeo, la hostilidad, y el falso apoyo. Para estos autores, existen dos fuentes principales de error en la entrevista: la que pertenece a la validez o significado de los datos y la que pertenece a su solidez o confiabilidad. Las tendencias para sacar conclusiones acerca de un individuo que se basan únicamente en la información de la primera impresión, limitan el significado y la precisión de los datos de la entrevista. Estas tendencias han sido denominadas como "*Efecto de Halo*" y "*Error general poco destacado*". Los malentendidos culturales pueden también sesgar los datos de la entrevista y llevar a conclusiones inexactas. Por otro lado, los coeficientes de validez predictiva para la información de la entrevista varían ampliamente. La confiabilidad de la información de la entrevista ha sido medida principalmente en términos del acuerdo entre entrevistadores respecto a variables tales como la inteligencia y los rasgos. Cuanto más estructurada sea la entrevista, mayor será el acuerdo entre los entrevistadores. De esta manera, al igual que los coeficientes de validez, los coeficientes de confiabilidad para la información de la entrevista varían ampliamente. La

capacitación tiende a mejorar la confiabilidad. Las habilidades para la entrevista se desarrollan mediante el conocimiento de una buena conducta de entrevista y sus principios, la práctica supervisada y un esfuerzo consciente por crear los hábitos correctos. Sin embargo, la entrevista es falible. La información de la entrevista puede ser mejor si se la considera como complemento de otras fuentes de información.

7.3. LOS AUTOINFORMES

Los autoinformes en sus distintas modalidades (inventarios, cuestionarios...) son considerados en la actualidad como una estrategia de gran importancia para la Evaluación Psicológica. En un principio, los autoinformes fueron rechazados por los evaluadores conductuales, sin embargo, a partir de los años setenta, se produjo una amplia utilización de este tipo de instrumentos en este enfoque.

7.3.1. DEFINICIÓN Y CLASES DE VARIABLES EVALUADAS MEDIANTE AUTOINFORMES

El autoinforme es definido como un mensaje verbal que un sujeto emite sobre cualquier tipo de manifestación propia (Fernández-Ballesteros, 2003, 2004d; Fernández-Ballesteros y Márquez, 2003). Un sujeto puede informar verbalmente sobre sus conductas motoras (número de cigarrillos fumados...), respuestas fisiológicas (presencia de taquicardias...), cogniciones (pensamientos, ideas de suicidio, atribuciones...), su experiencia subjetiva, sus atribuciones causales (del motivo de consulta...), como percibe una situación o a sus personas allegadas. Es decir, los autoinformes pueden referirse a múltiples manifestaciones tanto externas como internas.

El autoinforme puede ser considerado el método prioritario y directo por excelencia, a la hora de explorar los contenidos mentales del ser humano. Son instrumentos de gran utilidad para estudiar contenidos o productos mentales (productos de la actividad cognitiva, aquello de lo que el sujeto es consciente y puede verbalizar), resultan parcialmente útiles a la hora de evaluar procesos cognitivos (operaciones secuenciales necesarias para llegar a un determinado producto mental, por ejemplo, qué se hace para resolver un problema), y son irrelevantes a la hora de dar la razón de estructuras cognitivas (organizaciones estables y subyacentes a los fenómenos cognitivos).

Los autoinformes se han utilizado desde la Psicología Diferencial como base fundamental en la construcción de técnicas de medición de atributos intrapsíquicos a partir de los correlatos encontrados entre distintas manifestaciones verbales. También han sido utilizados para evaluar estados dependientes de situaciones concretas. Recientemente, desde el paradigma conductual-cognitivo los autoinformes están siendo utilizados en la recogida de información sobre las conductas problema.

Para Fernández-Ballesteros (1992d, 2004d) mediante los autoinformes se pueden explorar las siguientes clases de variables: rasgos o dimensiones de personalidad, estados, repertorios conductuales y repertorios cognitivos.

Rasgos o Dimensiones de la Personalidad: Los procedimientos de recogida de información más conocidos, los tests de personalidad (EPI, 16 PF, MMPI), son autoinformes tipificados que contruidos a través de procedimientos psicométricos, permiten obtener una

puntuación diferencial, es decir, la posición relativa del sujeto en una determinada variable intrapsíquica tras la comparación de sus respuestas con las del grupo normativo (perspectiva internalista).

Estados: Autoinformes que permiten explorar estados emocionales de ansiedad (STAI. Spielberger et al., 1970/1988), ira (STAXI-2, Spielberger, 2000), depresión (Spielberger, 2008)... del sujeto, que son dependientes de situaciones ambientales concretas (perspectiva situacionista).

Repertorios, procesos y estructuras cognitivas: Los autoinformes más utilizados para evaluar constructos cognitivos son: (1) autoinformes ligados a la percepción que el sujeto tiene de su ambiente (creencias, atribuciones...); (2) Automensajes o Autoinstrucciones que el sujeto se da ante la situación problema, y procesos o estrategias cognitivas que sigue en la realización de un problema; y (3) Funcionamiento motivacional (expectativas, estímulos internos o externos reforzantes...) (perspectiva conductual-cognitiva y cognitiva).

Repertorios Clínicos Conductuales: Autoinformes sobre la conducta motora, cognitiva y fisiológica, considerados como muestras de conducta, y no como indicador de rasgos intrapsíquicos. Autoinformes que tienen por objetivo evaluar distintos trastornos de conducta como miedos, depresión, obsesiones... (perspectiva conductual).

En el 2004, Fernández-Ballesteros incluye dentro de los autoinformes a la evaluación de construcciones idiográficas y narrativas, que en este proyecto docente se incluirán dentro de las técnicas subjetivas.

7.3.2. PRINCIPALES TIPOS DE AUTOINFORME

Fernandez-Ballesteros (2004d) clasifica los diferentes tipos de autoinformes en cuatro grandes apartados: 1) Entrevista - Autobiografía - Historia Clínica, 2) Cuestionarios - Inventarios - Escalas, 3) Autorregistros, y 4) Pensamientos en voz alta.

Entrevista, Autobiografía, Historia Clínica: La entrevista es la modalidad más extendida y el autoinforme más importante. Supone un intercambio, cara a cara, entre dos o más personas, una de las cuales requiere información y la otra se la brinda. Esta información puede referirse al propio sujeto, y en este sentido supone un autoinforme, o a otras personas o eventos relevantes. Los datos que se recogen con la entrevista pueden ser complementados con información proveniente de otros autoinformes como la autobiografía, y la historia clínica.

Cuestionarios, Inventarios y Escalas: Son autoinformes que presentan un listado de afirmaciones o preguntas a las que el sujeto debe: (a) responder si-no o verdadero-falso, (b) elegir u ordenar los elementos según las preferencias del sujeto, y (c) anotar el grado de conformidad según una escala ordinal o de intervalos (Fernández-Ballesteros, 1992d). Fueron muy utilizados en la Evaluación tradicional y actualmente asistimos a un gran desarrollo de los mismos. Se están readaptando instrumentos ampliamente contrastados y

se están generando otros nuevos (Ortíz, 1994). Una interesante recopilación y traducción de diversos cuestionarios, inventarios y escalas para evaluar ansiedad, depresión y habilidades sociales, ha sido realizada por Comeche, Díaz y Vallejo (1995), siendo este trabajo recomendado para la ampliación del tema de estudio. Muchos cuestionarios, inventarios y escalas no han sido estandarizados, mientras que otros han sido estandarizados pudiéndose incluir en el apartado de tests psicométricos.

Autorregistros: Se trata de un método de autoinforme consistente en un doble proceso: por un lado, el sujeto debe atender deliberadamente a la propia conducta y, por otro lado, registrarla mediante algún procedimiento establecido. Se pueden diferenciar 4 tipos de dispositivos de autorregistro: *Autorregistros de lápiz y papel* (el sujeto anota en un registro de papel la conducta objeto de estudio cuando aparece), *Contadores de respuesta* (el sujeto pulsa el botón cuando la conducta aparece), *Dispositivos de tiempo* (cronómetro para automedir la duración de la conducta objeto de estudio), y *Dispositivos electrónicos* (magnetofón en el que se graban pensamientos o informaciones sobre la variable objeto de evaluación).

Pensamientos en voz alta: Son autoinformes sobre conductas preferentemente cognitivas que se recogen en el momento de producirse; generalmente, en situaciones experimentales. Son técnicas no estructuradas ya que permiten cualquier tipo de respuesta, las verbalizaciones se registran, generalmente, en el laboratorio ya que responden a estímulos concretos suministrados por el evaluador. Se pueden identificar 3 procedimientos más frecuentes de recogida de los pensamientos en voz alta: (a) *Monólogo continuo* (el sujeto verbaliza sus pensamientos o sentimientos simultáneamente a la realización de una determinada tarea propuesta por el experimentador); (b) *Muestras de pensamientos* (el evaluador solicita en determinados momentos de la tarea que el sujeto verbalice lo que le pase por dentro); y (c) *Registro de eventos* (el sujeto señala cuándo ocurre una determinada actividad interna de interés para el examinador).

Esta es una posible clasificación de los autoinformes, pero como sucedía con las taxonomías de los instrumentos diagnósticos, se pueden adoptar diferentes criterios a la hora de operar una clasificación. Por ejemplo, Echeburua (1996c) diferencia dos tipos de escalas de evaluación conductual: (1) Según el contenido: Globales y Específicas; y (2) Según quien la cumplimente: autoaplicadas, heteroaplicadas o de observación, y analógicas visuales. Estas últimas consisten en una línea de longitud fija, habitualmente 10 cm, en cuyos extremos se sitúan los dos polos opuestos de la cuestión a medir, sin existir palabras que describan posiciones intermedias, debiendo la persona señalar en la línea el lugar que corresponde al estado percibido. Estas escalas, en opinión de Echeburua, son simples, pero eficaces cuando se trata de evaluar diariamente el estado de ánimo de una persona y comprobar si existen fluctuaciones, es decir, cuando se trata de medir los cambios de sintomatología reiteradamente en poco tiempo.

7.3.3. FUENTES DE ERROR DE LOS AUTOINFORMES

Los autoinformes producto de la introspección o la auto-observación han sido fuertemente criticados. Pese al enorme valor y desarrollo empírico de las técnicas de autoinforme en la actualidad siguen planteándose serias objeciones metodológicas. Se cuestiona, por un lado, su fiabilidad y validez, y por otro, la influencia que ciertas variables puedan tener sobre la calidad de los autoinformes. Varios trabajos (Cano y Miguel Tobal, 1994; Fernández-Ballesteros, 1993b) se han centrado en el estudio de la fiabilidad de distintos autoinformes.

Fernández-Ballesteros (1992d, 2003, 2004d) después de revisar las posibles distorsiones o fuentes de error de los autoinformes subraya 3 como las más importantes y frecuentes: Simulación, Deseabilidad Social y Tendencias de Respuesta.

- *La Simulación:* El sujeto evaluado falsea conscientemente sus respuestas. Existen autores que plantean que la capacidad de simulación es una variable de personalidad relacionada de forma directa con ajuste personal, mientras que otros consideran la simulación como un factor importante de error y han tratado de aminorar tal distorsión.
- *La deseabilidad social:* El individuo evaluado se describe involuntariamente de forma congruente con la imagen socialmente deseable. Muchos autores han puesto de relieve cómo una gran parte de la varianza en las respuestas a los autoinformes es explicada a través del afán de los sujetos de dar una imagen de sí mismos socialmente deseable.
- *Tendencias de Respuesta:* Las tendencias de respuesta están influidas por las alternativas planteadas en el autoinforme. Se destacan: (1) El asentimiento (tendencia que lleva al sujeto a responder sí o verdadero en los autoinformes de respuesta dicotómica, con independencia del contenido de la pregunta que se formule); y (2) Los errores escalares (tendencia a responder en el centro, o por el contrario, en los extremos de las escalas).

Zabalegui (1990) realizando un análisis de la problemática específica de los cuestionarios de personalidad subraya como sesgos:

- *Limitaciones por parte del examinado:* motivación, deseabilidad social, sinceridad, aquiescencia o tendencia a contestar en el sentido de la pregunta, e ignorancia o bien por falta de capacidad de introspección o por incomprendibilidad del cuestionario por bajo nivel cultural.
- *Limitaciones por parte del instrumento empleado:* validez, fiabilidad y objetividad.
- *Limitaciones de la situación de evaluación:* escalas de algunos instrumentos que son más situacionales mientras que otras son más independientes del momento presente.

Pese a todo, respecto a la valoración de los cuestionarios se puede afirmar que cuando han sido diseñados, valorados y aplicados de forma adecuada, pueden ser una fuente de datos útil y eficiente. No son costosos de administrar ni puntuar y tienen validez aparente para el cliente; su análisis e interpretación puede simplificarse mediante la administración y puntuación computerizada. Pero debido a la posibilidad de sesgos y a las dificultades inferenciales en la interpretación de puntuaciones derivadas, los datos deberían ser corroborados con datos procedentes de otros procedimientos de evaluación. Otras ventajas atribuidas al uso de

autoinformes es que pueden puntuarse de forma objetiva, minimizando la inferencia por parte del evaluador; además, son aplicables a un amplio rango de población y conductas-problema y pueden recoger información de los tres sistemas de respuesta.

7.4. LAS TÉCNICAS OBSERVACIONALES

7.4.1. DEFINICIÓN Y OBJETIVOS

Como ha señalado Fernández-Ballesteros (2004e), la observación es la estrategia fundamental del método científico, y, por tanto, cualquier técnica de recogida de información en evaluación psicológica conlleva, de una u otra forma, observación. La observación es, además, una conducta humana universal que ocurre, también, en la situación de evaluación entre el sujeto y el evaluador y que se produce sea cual fuere el modelo teórico de éste. Sin embargo, no sería correcto decir que todas las técnicas de recogida de información en evaluación psicológica son técnicas observacionales, existen una serie de características delimitadoras de lo que es un método observacional:

- Quien observa es un experto (participante o no participante en la situación) o ha sido entrenado para observar deliberadamente, aunque sea brevemente. Así, se incluye como método observacional la observación deliberada realizada por personas allegadas al sujeto (por ejemplo los padres que observan conductas de su hijo).
- Lo que se observa generalmente, es un hecho que ocurre en la situación de observación, preferentemente una situación natural o una situación artificial, preparada para reproducir el comportamiento que ocurre en la situación natural. Por supuesto, el “hecho” que se observa puede variar en su nivel de molaridad-molecularidad o la traducción que realiza el observador puede tener distintos niveles de inferencia. Por ejemplo, se pueden observar conductas (por ejemplo, si el niño moja la cama), se pueden inferir de las conductas atributos (por ejemplo, si el niño está nervioso) o se puede observar un hecho presente que representa la expresión de un conjunto de conductas pasadas (por ejemplo, cuando observamos la existencia o no de desechos). Sin embargo, algunos autores no incluyen como método observacional el que se expresa mediante atributos, sino sólo aquel referido a conductas motoras externas.
- Se observa de forma sistemática, por lo que la observación conlleva un protocolo u hoja de registro preparada al efecto que puede tener distintos niveles de estructuración, admitiéndose desde un protocolo simple que permite la anotación narrativa hasta los más sofisticados códigos de observación.

En definitiva, es observación cualquiera de los procedimientos y técnicas de recogida de información que requieren la percepción deliberada de comportamientos emitidos por un/os sujeto/s (a distintos niveles de complejidad e inferencia), realizada por un observador entrenado (participante o no participante) mediante protocolos preparados al efecto (de mayor o menor estructuración) que permitan una anotación sistemática, en una situación natural o artificial.

A través de la observación se pretende recoger mediante un registro objetivo, sistemático y específico, las conductas que se producen en un determinado contexto, sometiendo dicha información a una adecuada codificación y análisis (Anguera, 1988; Garaigordobil, 1999c).

La observación de la conducta en ambientes naturales es una parte importante del proceso de evaluación clínica y psicoeducativa, siendo una de las técnicas más utilizadas por los clínicos en la evaluación de problemas emocionales, sociales y de conducta de niños y adolescentes. Esto se debe a que en tanto que la observación directa de la conducta favorece la evaluación de las contingencias ambientales que están operando para producir, mantener o exacerbar la conducta del sujeto, su utilización se asocia a intervenciones ambientales más efectivas. La observación permite contextualizar el problema, recoger información relevante de cara a la previsión de cambios en el contexto, así como ayudar a modificar las variables que contribuyen a originar o mantener las dificultades del sujeto.

La observación conductual, en el contexto de la evaluación de niños y adolescentes, cubre las siguientes funciones dentro del proceso de evaluación (Sattler, 1996):

- Proporciona información sobre la conducta espontánea en el ambiente natural.
- Aporta información sobre la conducta interpersonal.
- Proporciona un registro sistemático de las conductas que pueden emplearse para evaluación y planes de intervención terapéutica.
- Permite verificar la precisión de los informes de padres y maestros acerca la conducta del niño.
- Posibilita la comparación entre la conducta en situación de prueba y en ambientes más naturales.
- Es muy útil en el estudio de niños pequeños y de aquellos con deficiencias en el desarrollo a los que resulta difícil evaluar mediante otros procedimientos.

7.4.2. UNIDADES DE OBSERVACIÓN Y DE MEDIDA

Existen distintas formas de enfocar el concepto de unidad de observación. Puede hacer referencia al objeto que se pretende estudiar (persona, grupo, institución...). Puede referirse a un concreto segmento de tiempo más o menos amplio del continuo de conducta. Así mismo, pueden tomarse como unidades de análisis las relaciones habidas entre las actividades de la/s persona/s y los elementos ambientales (estímulos físicos y sociales). Siguiendo la taxonomía de Fernández-Ballesteros (1992c, 2004e) y las observaciones de otros investigadores (Sechrest y Hill, 2003) se elabora el cuadro 24 en el que se considera unidad de observación aquel conjunto de eventos conductuales que con distintos niveles de molaridad-molecularidad son observados en Evaluación Psicológica.

Cuadro 24. Unidades de observación

UNIDADES DE OBSERVACIÓN (QUÉ OBSERVAR)	
Continuo del comportamiento	La observación se realiza sobre todo el continuo de conducta, tratando de registrar en forma descriptiva la mayor parte de los eventos que ocurren en un contexto natural en amplias unidades de tiempo. No se realiza una previa especificación de las conductas o atributos a observar. Se observa en tiempo real y forma continua. Las descripciones se realizan sobre aspectos verbales, no-verbales y/o espaciales de la conducta pudiendo ser compaginados estos con impresiones del observador sobre lo observado.
Atributos	Los psicólogos del rasgo y los psicodinámicos han utilizado la observación para la obtención de atributos o construcciones teóricas. Observan las conductas manifiestas (fruncimiento del ceño, temblor en las manos, salir de la habitación), y a través de un proceso de inferencia (preocupación, nerviosismo, evitación de la situación) se postula un rasgo (ansiedad) considerando que esta entidad está siendo expresada por la conducta manifiesta.
Conductas	Se observan conductas manifiestas motoras, verbales y fisiológicas. Ahora existe una previa especificación de la conducta que se va a observar. La formulación de estas unidades de análisis puede ser empírica o teórica. La observación de estas unidades permite mínimas inferencias por parte del observador. Se seleccionan los intervalos de tiempo en los que se va a realizar la observación y su duración.
Interacciones	El objeto de estudio es la relación funcional entre dos eventos que se producen secuencialmente procedentes de dos o más personas o entre una persona y una dimensión ambiental. Existe una previa especificación de las conductas a observar. Estas unidades están constituidas por influencias recíprocas existentes entre individuos o entre un individuo y un grupo o un determinado ambiente. Suele realizarse en unidades de tiempo previamente establecidas.
Productos de conducta	Una gran parte de la evaluación se realiza sobre productos de conducta, es decir, se observa el resultado de un conjunto de actividades internas o externas que él/los sujeto/s ha/n realizado en situaciones tanto naturales como artificiales. Pueden considerarse productos de conducta: <i>Medidas de erosión:</i> destrucciones, daños u otros cambios físicos que existen en el ambiente habitual del sujeto y que se han producido como producto de su conducta (marcas en el césped por el paso de los habitantes). <i>Medidas de huella:</i> productos de la conducta del sujeto al utilizar objetos o al consumirlos (botellas consumidas por un alcoholico). <i>Medidas de archivo:</i> Dibujos, notas del curriculum académico, documentos personales, clásicos tests de ejecución.

Anguera (2003ab) enfatiza que se puede observar: la conducta no verbal, la conducta espacial o proxémica, la conducta vocal o extralingüística y la conducta verbal lingüística, enfatizando las conductas manifiestamente observables, sin incluir los atributos o rasgos.

Todas estas unidades de observación pueden ser evaluadas en base a parámetros cuantitativos y cualitativos de las mismas, es decir, con distintas unidades de medida. Como subraya Carreras (1991/1993) independientemente de observar conductas, interacciones o atributos, a la hora de dar una versión cuantificada de las mismas habremos de proceder a la selección de las unidades de medida que van a ser adoptadas. Partimos de que el fenómeno observado se da en alguna medida, es decir, según una cierta cantidad en alguna de las propiedades o dimensiones de la unidad de análisis observada. En general, el tipo de medida dependerá de la unidad de observación, pudiéndose diferenciar las unidades de medida que se exponen en el cuadro 25.

Cuadro 25. Unidades de medida

UNIDADES DE MEDIDA	
Ocurrencia	Constatación de si un fenómeno se da o no.
Frecuencia	Número de veces que ocurre el fenómeno.
Duración	Tiempo que ocupa el fenómeno objeto de estudio.
Dimensiones cualitativas	<i>Intensidad o Magnitud</i> que se evalúa mediante el estudio del impacto de la conducta (intensidad de la agresión física, por el daño infligido al agredido). <i>Adecuación</i> evaluada mediante los aciertos o errores que el sujeto obtiene al ejecutar una tarea (medida utilizada principalmente en los tests de ejecución)...

7.4.3. TÉCNICAS DE REGISTRO

En términos generales, el registro es una transcripción de la representación de la realidad por parte del observador mediante la utilización de códigos determinados, y que se materializa en un soporte físico que garantiza su prevalencia (Anguera et al., 1991/1993). La estructuración del sistema de observación permite la replicabilidad y el control de los resultados. Pero hasta cierto punto la sistematización de la observación depende del marco referencial teórico del observador. Evaluadores partidarios de enfoques fenomenológicos y psicodinámicos suelen utilizar observaciones muy poco sistemáticas que permiten la descripción de los sujetos observados en función de una serie de impresiones o atributos de relevancia teórica. Por el contrario, los observadores conductuales y los que desde una perspectiva diferencialista utilizan productos de conducta emplean sistemas estructurados y estándar de observación en forma de códigos de categorías de conducta o de protocolos tipificados de recogida de datos (Fernández-Ballesteros, 1992c, 2004e).

Por otra parte, el nivel de estructuración puede depender también de la fase en la se efectúa la observación. Como señala Anguera (1981, 1983, 1989, 1990, 1994b, 2003ab), la observación ha de realizarse mediante un plan en el que en un primer momento se efectúa una observación escasamente sistematizada, con el fin de estudiar la forma de operativizar el evento a observar, fundamentalmente cuando no existen códigos tipificados sobre el fenómeno objeto de examen. En estos casos suele comenzarse con una observación no sistemática que permita establecer descripciones precisas de los eventos presentes, así como de las condiciones ambientales ante las que se producen, y que luego pasarán a ser observados en forma sistemática. Integrando las descripciones de varios autores (Anguera, 1994b; 2003ab;

Dieguez Sánchez, 1986; Fernández-Ballesteros, 1992c, 2004e, Hersen y Bellack, 2002, Fahrenberg, 2003) se ha elaborado el Cuadro 26 en el que se presentan varias técnicas de registro (con qué observar).

Cuadro 26. Técnicas de registro observacional

TÉCNICAS DE REGISTRO OBSERVACIONAL (CON QUÉ OBSERVAR)	
Registro Narrativo	El observador se limita a tomar nota escrita u oral (magnetofón) de lo que va ocurriendo en la situación de observación, sin que exista una estructuración previa que dirija o articule la recogida de datos y sigue un estilo textual. Pueden utilizarse como paso previo para el establecimiento de códigos estructurados.
Escalas de Apreciación	Un juez u observador que conoce al sujeto valora descripciones del sujeto mediante una escala de estimación. Son útiles para tener una primera aproximación cuantificada de las conductas problemáticas o adaptativas del sujeto, así como otras opiniones que tengan los allegados sobre él.
Catálogos de Conducta	También denominados listas de rasgos contienen una serie bien especificada de conductas, encuadradas o no en clases y con o sin indicación de antecedentes o consecuentes ambientales de las mismas. Los catálogos más frecuentes son los registros de conductas y las matrices de interacción.
Códigos o Sistemas de Categorías	Conllevan la enumeración, descripción y clasificación de los eventos conductuales y/o contextuales que se pretenden observar así como el modo de llevar a cabo la observación. Son los registros más sofisticados y comportan un largo proceso y sofisticadas elaboraciones para su construcción. Para una clarificación del procedimiento de construcción de un sistema de categorías consultar Anguera (1991/1993, 1944b, 2003ab). Existen códigos estandarizados y tipificados presentando garantías para su utilización. Para una revisión de los códigos de categorías disponibles consultar los catálogos de instrumentos de evaluación de Hersen y Bellack (1988, 2002), y TEA (2009). Estos códigos proveen información sobre conductas y/o interacciones complejas. Además, al ser un instrumento estándar presentan mayores garantías científicas.
Registro de Productos de Conducta	Se han elaborado complejos sistemas de registro y puntuación sobre productos de conducta. El sujeto realiza una tarea determinada (rompecabezas, historia, dibujo....) y el evaluador puntúa sus ejecuciones según unas dimensiones de respuesta (tiempo, aciertos, errores...). Así, los resultados de los tests de inteligencia, aptitudes, rendimiento, habilidades lingüísticas... podrían ser conceptualizados como registros del comportamiento final de un sujeto en una situación estándar.
Procedimientos Automáticos de Registro	Dentro de tales procedimientos pueden incluirse tres grupos: • <i>Medios técnicos de registro auxiliares del observador:</i> Aparatos como el Datamyte 900, el Sistema 7 (SSR) o el Sistema de puntuación de observación conductual (BOSS)... que permiten registrar automáticamente eventos de conducta a través de categorías previamente establecidas que pueden ser almacenadas en una

	computadora pudiéndose obtener información sobre su ocurrencia, frecuencia, duración, secuencia.... • <i>Aparatos de registro a distancia ocultos</i>: Procedimientos como el magnetofón, el vídeo-tape o los espejos unidireccionales son utilizados con mayor frecuencia, pero por razones éticas su empleo requiere la autorización del sujeto observado. • <i>Observación mediante aparatos</i>: Registros psicofisiológicos, dispositivos de registro de conducta motora y los aparatos de medida de productos de conducta, considerados técnicas objetivas.

7.4.4. CONSIDERACIONES MUESTRALES

Las limitaciones y dificultades de la observación han llevado a la necesidad de utilizar procedimientos de muestreo. El objetivo de tales procedimientos es el de obtener muestras representativas de los eventos objeto de estudio (Quera, 1991/1993). Cuando un evaluador decide utilizar la observación debe de plantearse (Fernández-Ballesteros, 1992c, 2004e):

- Durante cuanto tiempo se va a prolongar la observación.
- Con qué frecuencia va a observarse.
- En qué momentos se van a iniciar y terminar los períodos de observación y si éstos van a ser constantes a lo largo de los períodos de observación o van a variar en cada unidad de observación.
- Si se van a utilizar intervalos de tiempo para la observación y el registro dividiendo así los períodos de observación.
- Si se pretende tener constancia de lo que ocurre en distintas situaciones, habrá que decidir en qué situación se va a realizar la observación.
- Si se trata de observar a un sujeto o a varios habrá de seleccionarse a qué sujeto y en qué momento el observador registrará la conducta objeto de estudio.

Todas estas decisiones implican diferentes tipos de muestreo como son *muestreo de tiempo* (se decide durante cuanto tiempo va a observarse, con qué frecuencia y en qué momentos), *muestreo de situaciones* (se explora la conducta objeto de estudio en distintas situaciones), o *muestreo de sujetos* (se decide el orden y los intervalos de observación de los sujetos de un grupo).

Para elegir una técnica de muestreo el investigador deberá considerar cuidadosamente las características conductuales y de los sujetos observados, así como las preguntas que se formularon al inicio del estudio. Desde esta consideración Anguera (1994b), plantea cuatro tipos de muestreo, el *muestreo ad libitum* (corresponde a experiencias no sistematizadas y se trata de notas de campo obtenidas con el único criterio de juzgarse interesantes para quien las recoge, o bien registrar todo lo que se puede o lo que es más fácilmente observable), *muestreo de eventos* (selección como unidades de la muestra de todas las ocurrencias de una conducta o gama de ellas que tengan lugar a lo largo de las sesiones establecidas), *muestreo focal* (se registran todas las acciones de un sujeto previamente seleccionado, y las que se le dirigen), y *muestreo de tiempo* (que se puede plantear en dos modalidades: muestreo instantáneo o de puntos de tiempo, y muestreo de intervalos).

7.4.5. FUENTES DE ERROR Y REQUISITOS CIENTÍFICOS DE LA OBSERVACIÓN

Han sido varios los autores que analizando la calidad de la observación (Anguera, 1983, 1988, 1994b, 2003ab; Dieguez Sánchez, 1986; Fernández-Ballesteros, 1983, 1992c, 2004e; Haynes y O'Brien, 2000; Hartman, Barrios y Wood, 2003; Ibañez, 1993a; Martorell, 1988d) que han señalado diversas fuentes de error de la metodología observacional y sus técnicas. En conjunto han subrayado tres fuentes de error: Sesgos procedentes del sujeto observado, del observador, y del sistema de observación utilizado. Integrando las aportaciones de estos autores se ha elaborado el Cuadro 27 con los sesgos señalados con mayor frecuencia.

Cuadro 27. Fuentes de error de la observación

FUENTES DE ERROR DE LA OBSERVACIÓN	
Del Sujeto Observado	• <i>Reactividad</i>: Modificación de la conducta del sujeto observado por el solo hecho de saberse observado. • <i>Autorreactividad</i>: Es la influencia que ejerce el autorregistro sobre la ocurrencia de conducta, siendo el único caso en que debe valorarse la reactividad como efecto positivo ya que puede ejercer un efecto de autocontrol en un proceso terapéutico. • <i>Características de los sujetos observados</i> (edad, competencias, sexo...).
Del Observador	• <i>Grado de Participación del observador</i>: A mayor grado de participación o interacción entre observador y observado mayor fuente de influencia. • <i>Expectativas</i>: Las expectativas que el evaluador lleva a la situación influirán en la misma (Rosenthal, 1966/1976), pudiendo introducir sesgos al elegir el sistema de observación, al registrar las conductas o categorías conductuales consonantes con las hipótesis previas, o al seleccionar el diseño estadístico de elaboración de los datos procedentes de lo observado. • <i>Características Personales</i>: Las características o atributos del observador (sexo, edad, habilidades...) pueden también mediar en los resultados de la observación. • <i>Reactividad recíproca</i>: Tiene lugar cuando el sesgo de reactividad afecta también al observador, que si es a la vez el evaluador, se ve influenciado al saber que el sujeto observado no actúa espontáneamente, dado que se siente protagonista de la situación de observación. • <i>Entrenamiento</i>: El nivel de entrenamiento influye en la calidad de la información registrada, siendo este entrenamiento esencial para evitar diversas fuentes de sesgo.
Del Sistema de Observación	• <i>El tipo de registro</i>: El registro elegido afecta a la fiabilidad y validez de los datos. Cuando las definiciones conductuales son claras, se utiliza un número reducido de categorías, o se utilizan códigos estándar, se incrementa la bondad de los datos recogidos. • <i>El tipo de muestreo</i>: El tipo de muestreo utilizado también será una fuente de influencia. • <i>Características de los aparatos de registro. Procedimientos eléctricos o mecánicos de registro de los datos observacionales</i>: Estos procedimientos incrementan la objetividad y la exactitud de la observación, sin embargo, algunos de ellos plantean problemas metodológicos que han de ser tenidos en cuenta antes de su utilización.

Además de estas tres fuentes de sesgo, Anguera (1988) llamó la atención sobre otro grupo de problemas que se pueden dar en la metodología observacional, denominados **fallos procedimentales**. Estos fallos pueden ser de carácter diverso y referirse a distintas fases del proceso de observación. Estos fallos serían: *Punto de mira incorrecto*; *fallos de omisión*, es decir, no tener en cuenta la ocurrencia de una conducta; *fallos de comisión* o registro de una conducta distinta a la ocurrida; *inadecuada planificación del muestreo*; *fallos del funcionamiento del material de registro*; *mala definición de las categorías*; y *falta de sincronización entre observadores* de un equipo respecto a los inicios y finales de sesión... Los datos observacionales pueden estar sometidos a diferentes sesgos, por lo que el investigador debe contrastar la fiabilidad, validez y exactitud de los mismos.

En lo que respecta a la **fiabilidad**, de los diversos tipos de fiabilidad existentes, la más relevante es la *fiabilidad interjueces*, es decir, dos observadores independientes observando las mismas conductas y con el mismo protocolo de observación coinciden o no en sus registros. El concepto denominado desde la psicometría clásica como fiabilidad interjueces ha sido reformulado desde la teoría de la generalizabilidad haciendo referencia al universo de los observadores (Fernández-Ballesteros, 1992c; Fernández-Ballesteros y Calero 2004). El trabajo de Pons y Frías (1995) proporciona una guía para que los investigadores usen e interpreten adecuadamente los niveles de fiabilidad hallados en los trabajos que emplean procedimientos de registro observacional.

En la revisión realizada por Ibañez (1993a) sobre los procedimientos normalmente empleados para calcular la fiabilidad interjueces se destacan dos tipos de métodos: (1) Métodos para sistemas de registro de intervalo, como son el porcentaje de acuerdo y el estadístico Kappa; y (2) Métodos para el total de la sesión, entre los que destacan dos métodos como son la correlación de Pearson y el análisis de varianza. Debido a la importancia de la fiabilidad de los datos y puesto que los resultados pueden diferenciarse en función del índice de acuerdo utilizado, es conveniente que conste el procedimiento utilizado para su cálculo, el número de sesiones utilizadas, el índice aplicado, el tipo de registro y los criterios de acuerdo utilizados en la confrontación de los registros (Navas, 2001).

En lo que respecta a la **validez**, y a pesar de la importancia de este concepto, es curioso observar cómo la validez de los datos observacionales suele asumirse y no demostrarse. La validez de estos datos debe determinarse por la relevancia y significatividad de la conducta observada (Sattler, 1988, 1992, 1996). Bernstein y Nietzel (1982) ejemplifican cómo tratar este problema. Empleando la metodología observacional para estudiar la conducta hiperactiva, formulan las siguientes preguntas: (1) ¿Los códigos conductuales constituyen una definición funcional y satisfactoria de hiperactividad? (validez de constructo); (2) ¿Los datos reflejan la naturaleza y grado de la conducta observada? (validez de contenido); (3) ¿Los códigos conductuales predicen otros criterios importantes? (validez predictiva); y (4) ¿La conducta de la persona bajo observación refleja exactamente sus respuestas en otras situaciones? (validez concurrente). A pesar de que en esas 4 preguntas se pueden encontrar elementos de cierta arbitrariedad (la dificultad de diferenciar la validez de contenido, de la de constructo, por ejemplo), supone un buen comienzo para analizar la validez de los datos de observación (Ibañez, 1993a).

La validez de los datos de observación está íntimamente relacionada con la relevancia de las categorías de observación seleccionadas y con la exhaustividad del sistema, ya que el mismo debe incluir de forma completa todos los aspectos necesarios para operativizar el

constructo psicológico o conducta estudiada. Las principales amenazas a la validez de las observaciones conductuales son según Merrell (1999): a) categorías de observación pobremente definidas, b) la falta de datos de comparación social, c) la reactividad del observador, d) la especificidad situacional de la conducta, e) sistemas de registro inadecuados, y f) sesgos debidos a las expectativas del observador.

El análisis realizado por Fernández-Ballesteros (1992c, 2004e) en relación a las garantías científicas de la observación, y en concreto en relación a las garantías que deben presentar los datos recogidos con o sin procedimientos estándar, le lleva a plantear la necesidad de dar respuesta a 4 preguntas que hacen referencia a la generalizabilidad de las puntuaciones, tiempo y situaciones: (1) ¿En qué medida son generalizables los datos procedentes de un observador a los recogidos por otros observadores? (es decir, generalizabilidad del universo de observadores o fiabilidad interjueces); (2) ¿Hasta qué punto lo observado en un determinado momento es extensible a otros momentos de la vida del sujeto? (es decir, generalizabilidad del universo temporal o fiabilidad test-retest); (3) ¿Hasta qué punto los datos de observación procedentes de una situación son generalizables a otras situaciones? (es decir, generalizabilidad intersituacional o validez ecológica); y (4) ¿Hasta qué punto los datos evalúan el constructo que pretendemos observar?. Algunos investigadores (Haynes y O'Brien, 2000) subrayan que las garantías que ha demostrado la observación están mucho más en relación con la fiabilidad y el universo de generalizabilidad de los observadores que con la validez, sea de contenido, criterial o de constructo.

En relación a la validez de la observación, Haynes (2001), reflexiona sobre la observación conductual análoga que se refiere a la evaluación de la conducta evidenciada por un cliente en una situación controlada similar a las situaciones que probablemente encuentre en su entorno natural. Se trata de un tipo de observación que resulta especialmente útil en el ámbito clínico. Sin embargo, para Haynes (2001), los instrumentos para la observación conductual análoga a menudo se desarrollan sin prestar suficiente atención a sus características psicométricas, especialmente en lo que se refiere a su validez de contenido. Si bien dichos instrumentos pueden ser sensibles al cambio, su validez puede deteriorarse con el tiempo, y resulta afectada por numerosas fuentes de varianza. Por ello, tanto los investigadores que desarrollan instrumentos para la observación conductual análoga como los clínicos que los están utilizando deberían: a) especificar las decisiones que van a fundamentarse en las medidas de los instrumentos; b) especificar las conductas, relaciones funcionales, dimensiones, modelos de respuesta, y constructos que van a medirse con los instrumentos; c) seleccionar escenarios, situaciones, e instrucciones que maximicen su validez ecológica; y d) ser conscientes de que la validez de los mismos puede variar en función de las diferencias individuales.

Según Fernández-Ballesteros (2004e), tal vez la cuestión más importante esté no tanto en la validez del instrumento de observación utilizado como en las inferencias que puedan realizarse desde esas observaciones. En nuestro caso, y dado que nos estamos refiriendo al caso individual, lo importante a la hora de asegurarnos de la validez de nuestras inferencias sobre ese caso en particular es no realizar ninguna sin haber utilizado distintos métodos. Así, los resultados procedentes de los métodos observacionales deberán ser contrastados (en nuestro caso particular) con lo que hemos obtenido mediante otros métodos (entrevistas, cuestionarios, autorregistros...). Esa será la mejor manera de avalar nuestras inferencias. La autora enfatiza la existencia de paquetes estadísticos para el análisis de los datos observacionales, que pueden ser encontrados en Internet en las siguientes páginas:

www.efdeportes.com/efd20/trans.htm (Transcriptor). www.skware.com (BEST). www.efdeportes.com/efd18/codex.htm (Codex). www.psysoft.com/(Observe) . www.ub.es/comporta/comporta.htm (SDISGSEQ). http://www.noldus.com (Tbe Observer).

A modo de síntesis, se puede subrayar la importancia de planificar a priori la observación, teniendo en cuenta las fuentes posibles de sesgo y los mecanismos a utilizar a fin de neutralizarlos. Diversos autores como Fernández-Ballesteros (1992c, 2004e), o Ibañez (1993a) plantean guías de ayuda para facilitar la sistematización del procedimiento. De la integración de las observaciones de ambos se desprende que en la planificación de una observación deben ser clarificados los siguientes interrogantes:

- ¿Requiere el caso observación? ¿Propósito de la observación?
- ¿Qué unidades de observación están implícitas en la hipótesis sobre el caso? (continuo de comportamiento, atributos, conductas, interacciones, productos de conducta...).
- ¿Características y tipo de variable objeto de observación?
- ¿Qué unidades de medida van a ser utilizadas? (frecuencia, ocurrencia, duración...).
- ¿Mediante qué técnicas van a ser recogidas las observaciones? (registros narrativos, escalas de apreciación, catálogos de conducta...) ¿Calidad de los datos?
- ¿Quién o quiénes van a ser observados?
- ¿Dónde se va a observar? (situaciones naturales o artificiales).
- ¿Durante cuánto tiempo se realizará la observación? (tiempo total de observación, número de sesiones, duración de las sesiones, períodos de habituación...).
- ¿Quién o quiénes van a observar? (procedimiento de formación de los observadores).
- ¿Medios disponibles?
- ¿Cómo van a apreciarse la bondad de las observaciones? (fiabilidad y validez).

El procedimiento que conforma la metodología observacional (Anguera, 2003a) consta de una serie de fases lógicas:

- Fase exploratoria.
- Disposición previa (mantenimiento de la constancia intersesional, mantenimiento de la constancia intrasesional, tratamiento de las disrupciones temporales, temporalización e identificación de la sesión de observación).
- Plan de muestreo observacional.
- Elaboración del instrumento (sistema de categorías, formatos de campo, combinación de formatos de campo y sistemas de categorías, escalas de estimación o apreciación).
- Registro y codificación.
- Métrica del registro u obtención de parámetros.
- Control de calidad del dato.
- Análisis de datos.

Esta metodología además enfatiza la importancia de llevar a cabo una buena formación de los observadores que posibilite una observación con garantías de calidad.

7.5. TÉCNICAS PROYECTIVAS

7.5.1. DEFINICIÓN E HITOS HISTÓRICOS DE LAS TÉCNICAS PROYECTIVAS

Las técnicas proyectivas han sido y siguen siendo instrumentos muy utilizados en evaluación tanto en psicología clínica, en psicología escolar, como en otros campos de aplicación de la evaluación psicológica como es la psicología de la salud (Aguirre, 1992; Aguirre, Alamar y Pérez, 1995; Avila Espada, 1986, 1990; Fernández-Ballesteros, Vizcarro y Márquez, 1992; Marqués, Vizcarro y Fernández-Ballesteros, 2004; Meyer y Archer, 2001).

Fue Frank quien, en 1939, acuñó la expresión técnicas proyectivas para referirse a aquellos métodos de evaluación en los cuales el sujeto evaluado emite libremente respuestas a estímulos poco estructurados (del tipo de manchas de tinta, láminas o frases incompletas). Técnicas que se basan en el supuesto de que las personas proyectan sobre la realidad exterior sus percepciones desde el propio psiquismo (Aiken, 1996; Anastasi y Urbina, 1998; Avila, 1997; Gregori, 2001; Kaplan y Saccuzzo, 2006).

Estas técnicas han sido definidas por Lindzey (1961) como aquellos instrumentos considerados especialmente sensibles para revelar aspectos inconscientes de la persona que provocan una amplia variedad de respuestas subjetivas; son altamente multidimensionales y evocadores de datos inusualmente ricos con un mínimo de conocimiento por parte del sujeto del objetivo del test. Generalmente el material estimular presentado es ambiguo, estimula la imaginación del sujeto, de ellos no se derivan respuestas correctas e incorrectas y su interpretación depende de un análisis global. De un modo más genérico, el *Educational Testing Service* (ETS) (1991) consideró a las técnicas proyectivas como aquellas medidas que requieren una respuesta no estructurada en orden a efectuar una evaluación de la personalidad o cognición del sujeto examinado. Actualmente Silva (2003) designa con estos términos al conjunto de instrumentos cuyo objetivo es describir y caracterizar la personalidad.

Estas técnicas han sido conceptualizadas principalmente a partir del modelo psicoanalítico, siendo procedimientos de recogida de información útiles desde el enfoque dinámico puesto que sirven en el análisis del mundo inconsciente del sujeto. Para Fiske (1971) las técnicas proyectivas han de ser consideradas como procedimientos de observación, a través de un material tipificado, en las que el evaluador elicitaba una variedad de respuestas que se supone son expresión del mundo interno del sujeto.

Los tres hitos históricos fundamentales de estas técnicas, más frecuentemente citados, son:

- El establecimiento de la técnica de Asociación de Palabras de Jung en 1910.
- La primera edición del Psychodiagnostik de Rorschach en 1921.
- La publicación de Murray *Explorations in Personality*, base del Test de Apercepción Temática, en 1938.

Durante más de cuarenta años las técnicas proyectivas tuvieron un gran apogeo y desarrollo, siendo importantes exponentes de la Evaluación Psicológica. Pese a que a partir de los años sesenta comienzan a producirse abundantes críticas a estas técnicas, debido a resultados negativos sobre fiabilidad y validez de las mismas, se han seguido utilizando. A partir de los 80 existe una clara recuperación de estas técnicas y durante los 90 han ocupado un papel importante en los programas de Evaluación Psicológica y/o Psicodiagnóstico tanto en

España como en EEUU (Fernández-Ballesteros, 1991c), y en cuanto a los programas académicos, la presencia de las técnicas proyectivas ha sido constante, tanto en España como en EEUU (Marqu  ez et al., 2004).

A pesar de que han sido desarrolladas, en su mayor  a, desde un marco conceptual psicodin  mico, tambi  n han sido conceptualizadas desde el enfoque de la psicolog  a de la Gestalt, y espec  ficamente desde aquel que vincula la personalidad a la percepci  n. As   mismo, tambi  n el paradigma E-R y, concretamente, el modelo mediacional de Hull ha sido utilizado como soporte de las t  cnicas proyectivas.

La imposibilidad de encontrar una   nica teor  a que permita conceptualizar las t  cnicas proyectivas ha provocado que el grueso de investigaci  n se haya dirigido a la obtenci  n de pruebas correlacionales, en su mayor parte cl  nicas, que han ido alej  ndolas de la psicolog  a experimental y acerc  ndolas simult  neamente al modelo psicoanal  tico hasta el punto de convertirse en los instrumentos prioritarios de tal enfoque (Fern  ndez-Ballesteros et al., 1992).

7.5.2. CARACTER  STICAS DE LAS T  CNICAS PROYECTIVAS

Como subrayan Marqu  ez et al. (2004), independientemente del enfoque te  rico desde el que se pretenda explicar estas t  cnicas, lo que est   claro es que los psic  logos proyectivos suelen admitir la existencia de una serie de supuestos b  sicos en ellas, como son:

- Se parte de que el sujeto cuenta con una estructura b  sica y estable de personalidad. Tal estructura est   integrada por ciertas dimensiones, rasgos o construcciones, organizadas en forma idiosincr  sica en cada sujeto. Las respuestas de los sujetos ante estas t  cnicas permitir   la exploraci  n de esta estructura de personalidad.
- La organizaci  n de la estructura de la personalidad hace necesarios distintos niveles de profundidad en su an  lisis. As  , mientras algunas t  cnicas como el Rorschach permiten ahondar en los aspectos estructurales de la personalidad, otras como el TAT, analizar  n aspectos m  s superficiales y, por tanto, influenciab  les por la situaci  n.
- Existe una relaci  n entre los inobservables que integran la estructura de la personalidad y las manifestaciones conductuales de los sujetos, por lo que el an  lisis de la estructura de la personalidad, a trav  s de estas manifestaciones, permitir   la predicci  n del comportamiento.
- Toda respuesta ante el material proyectivo no es casual sino que es significativa y ser   entendida como un signo de la personalidad del sujeto.
- Cuanto m  s ambiguas sean las propiedades de los est  mulos de una t  cnica proyectiva tanto m  s reflejar  n las respuestas su personalidad.
- El sujeto no es consciente de la relaci  n entre sus respuestas y su mundo interno y es dif  cil que pueda falsear sus respuestas, por lo que estas t  cnicas pueden ser consideradas como enmascaradas.
- El an  lisis al que son sometidas las respuestas de los sujetos a las t  cnicas proyectivas ha de ser fundamentalmente cualitativo y global.

Entre las caracter  sticas de estas t  cnicas se ha subrayado el hecho de que presentan est  mulos ambiguos, requieren respuestas de la imaginaci  n, y son t  cnicas enmascaradas. A pesar de estos supuestos b  sicos caracter  sticos,   stas t  cnicas no son un grupo homog  neo

de procedimientos, ya que tanto los estímulos presentados al sujeto, como las respuestas requeridas difieren enormemente. Estos elementos diferenciales de las mismas han constituido ejes de clasificación de este tipo de técnicas.

7.5.3. EL CONCEPTO DE PROYECCIÓN

La hipótesis proyectiva sostiene que un individuo da estructura a estímulos que no están estructurados de una manera consistente con el propio patrón único de necesidades, temores, deseos, impulsos, conflictos y formas de percibir y responder conscientes e inconscientes (Cohen y Swerdlik, 2001). De acuerdo con la hipótesis proyectiva, las interpretaciones de un estímulo ambiguo o vago reflejan las necesidades propias de un sujeto, sus sentimientos, experiencias, condicionamiento previo y procesos de pensamiento entre otros (Kaplan y Saccuzzo, 2006). Avila Espada (1986, 1990, 1997) ha revisado distintas hipótesis del concepto de proyección destacando varias hipótesis desde las que se entiende este concepto de distinto modo:

1. **Hipótesis Proyectiva Clásica:** Emanada de concepto de proyección inicialmente propuesto por Freud con defensa primaria, según la cual el sujeto percibe en función de los propios conflictos y que encuentra expresión patológica en los trastornos de tipo paranoide.
2. **Hipótesis Holística-Funcional:** Caracterizada por concebir la personalidad como un “todo” en interacción dinámica con el entorno físico y socio-cultural. Esta hipótesis dio origen a un nuevo enfoque de las técnicas proyectivas conocido como “percepción-personalidad”, en el que se llegó a sustituir la nominación de proyectivas por los términos de apercepción y distorsión aperceptiva.
3. **Hipótesis de la Desviación:** Desde esta hipótesis se plantea que las técnicas proyectivas pueden ser vistas como el intento de estudiar las desviaciones personales de las pautas de respuesta que emergen como elementos comunes de los mundos personales de la mayoría de la gente.
4. **Hipótesis del Nivel de Adaptación:** Según esta hipótesis la gran ambigüedad de los estímulos presentados en estas técnicas incrementa la influencia de la diferencias individuales en la percepción. Las respuestas serán consideradas según tres formas de percepción del estímulo: como foco principal, como trasfondo del campo perceptivo, o como residuo de experiencias anteriores. Estas formas de percepción se asocian con niveles de adaptación del sujeto.
5. **Hipótesis de los Estados Directivos:** Basada en la teoría según la cual el proceso perceptivo está dirigido y controlado por ciertos estados específicos del organismo del individuo (necesidades básicas, valores, experiencia, variables de personalidad...).
6. **Hipótesis de la Actividad Perceptiva:** El proceso perceptivo se considera como un proceso activado por la disposición.
7. **Hipótesis de la Situación Experimental:** Supone el intento de adaptar las Técnicas Proyectivas al paradigma E-O-R, e incluso al paradigma E-R. Se distinguen tres tipos de situaciones para explicar las relaciones entre estas técnicas y los hábitos: situación de origen, situación de test, y situación de criterio. La relación entre estas tres situaciones se debe a los hábitos emocionales que fueron aprendidos en la situación de origen, permanecen en el test y permiten predecir el comportamiento en la situación criterio, dado su poder explicativo de la conducta. Según esta hipótesis las Técnicas Proyectivas se aplicarían

en el contexto de una situación experimental en la cual el control de determinadas variables posibilita el análisis de las respuestas producidas.

- 8. Hipótesis Cognitivo-Experimental:** Atiende a los contenidos y a los modos de procesamiento de la información. Desde esta hipótesis se aplica un modelo de procesamiento simple a la situación proyectiva, intentado a través de la manipulación de las dimensiones estímulares, lograr esclarecer lo que ocurre desde la presentación del estímulo hasta que se produce la respuesta.
- 9. Hipótesis Cognitiva-Aplicada:** Desde esta hipótesis la proyección es entendida como la manifestación de conductas encubiertas del individuo, en relación a estímulos relevantes. Se considera que la respuesta verbal de un sujeto ante una situación interpersonal definida, pero inacabada, pone de manifiesto su capacidad de resolver ese tipo de problemas y que tal capacidad cognitiva puede ser un buen predictor del ajuste de la persona en la vida real.
- 10. Hipótesis Existencial-Humanista:** Perspectiva sistematizada en los 80 que surge de los enfoques humanistas y existenciales sobre la personalidad, la psicoterapia y la clínica. Desde esta hipótesis se considera que las técnicas proyectivas nos ayudan a conceptualizar la experiencia de las demás personas, sistemática y consistentemente, y para comunicar nuestro entendimiento de sus experiencias.

7.5.4. CLASIFICACIÓN DE LAS TÉCNICAS PROYECTIVAS

En cuanto a las clasificaciones de las Técnicas proyectivas, es obligado, remitirse, en primer lugar, a la expuesta por Frank (1939) que consideró 5 categorías:

- Técnicas proyectivas constitutivas: El individuo evaluado constituye una estructura sobre un material que en sí no la posee (Test de Rorschach).
- Técnicas proyectivas constructivas: Presentan un material prediseñado (Dibujo de la Figura Humana de Machover, Test de los mosaicos).
- Técnicas proyectivas interpretativas: El sujeto debe ofrecer significados a diversos estímulos o situaciones estímulares (Test de Apercepción Temática de Murray).
- Técnicas proyectivas catárticas: Se incluyen en esta categoría aquellas que propician la catarsis (El juego).
- Técnicas proyectivas refractarias: Se trata de instrumentos de carácter expresivo (Tests grafológicos).

Lindzey (1959) estableció seis criterios de clasificación según: (1) los atributos del material propuesto; (2) las técnicas empleadas en su construcción (racional / empírica); (3) la forma de interpretar el test (análisis formal/contenido); (4) el propósito del test; (5) la forma de administración (colectiva/individual); y (6) el tipo de respuesta elicitada.

Posteriormente, Anastasi y Urbina (1998) las han categorizado en: 1) técnicas de las manchas de tinta (Rorschach, Hotzman...), técnicas pictóricas o gráficas (TAT de Murray, CAT de Bellack, Frustración de Rosenzweig...), técnicas verbales (asociación de palabras de Jung...), recuerdos autobiográficos (autobiografías, recuerdos tempranos) y técnicas de ejecución que incluyen tanto técnicas de dibujo como de juego (dibujo de la figura humana de Machover, dibujo de la familia de Corman, la casa, el árbol y la persona de Buck, técnicas de

juego..). Desde otra perspectiva, Fernández-Ballesteros (1980) y Marquéz et al. (2004) plantean 5 categorías de clasificación, tomando como referente la tarea que debe realizar el sujeto:

- Técnicas estructurales: Se presenta material visual de escasa estructuración que el sujeto debe estructurar diciendo lo que ve, o a lo que se puede parecer (por ejemplo, el Rorschach).
- Técnicas temáticas: Se presenta material visual con distintos grados de estructuración formal de contenido humano o parahumano sobre los cuales el sujeto debe narrar una historia (por ejemplo, el Test de Apercepción Temática de Murray, o el Test de Relaciones Objetales de Phillipson).
- Técnicas expresivas: Se da al sujeto la consigna verbal o escrita de dibujar una/s figura/s (por ejemplo, el Test de la Figura Humana, el Test del Árbol, el Test de la Familia...).
- Técnicas constructivas: Se aporta material concreto que el sujeto debe de organizar en base a distintas consignas (por ejemplo, test del pueblo).
- Técnicas asociativas: Se aporta una consigna verbal o escrita por la que el sujeto ha de manifestar verbalmente sus asociaciones frente a palabras, frases o cuentos (por ejemplo, las Fábulas de Duss, o el Test de las Frases Incompletas).

Recientemente, Valdés y Gómez de Terreros (2003) las han categorizado en técnicas estructurales (Rorschach), temáticas (TAT, CAT, FTT, Duss...) y expresivas (dibujo de la casa, el árbol, la figura humana, la familia...). Con la finalidad de identificar algunos de los tests más significativos y de mayor utilización incluidos en cada categoría, se ha elaborado el cuadro 28.

Cuadro 28. Tipos de técnicas proyectivas

TÉCNICAS	EJEMPLOS Y REFERENCIAS BIBLIOGRÁFICAS
Estructurales	• Psicodiagnóstico de Rorschach, 1921. Rorschach, 1921/1972; Exner, 1978, 1990, 1991, 1995.
Temáticas	• Test de Apercepción Temática (TAT) de Murray, 1938. Murray, 1938; Avila Espada, 1986b. • Test de Apercepción Temática (CAT) de Bellak 1949-1966. Bellak, 1966, 1979; Baringoltz de Hirsch, 1974. • Test de Relaciones Objetales de Phillipson (TRO) de Phillipson, 1955/1965; Stricker y Healey, 1990. • FTT. Tests de los Cuentos de Hadas. Coulacoglou (1995).
Expresivas	• Dibujo de la Figura Humana de Machover, 1949. Machover, 1949. Koppitz, 1976. Maganto y Garaigordobil, 2009. • HTP. Casa-Árbol-Persona de Buck, 1949. Hammer, 1969/1982. Buck y Warren, 1995. • El test del dibujo de la Familia de Corman, 1967. Corman, 1961/1967; Frank de Verthelyi, 1985.
Constructivas	• El juego del mundo y el test de la aldea imaginaria de Mucchielli, 1964. • Caja de juego diagnóstico-terapéutico. Efron y otros, 1986; Silva, 1991.
Asociativas	• Test de Asociación de palabras de Jung, 1910. Rapaport, 1965. • Frases incompletas de Kelly y Fisher, 1951. • Fabulas de Duss, 1950; Duss, 1950/1980; Miller y otros, 1989.

Entre las pruebas proyectivas de mayor utilización y que gozan de amplia investigación cabe destacar la prueba del test de las manchas de tinta de Rorschach y el Test de Apercepción Temática de Murray (Aiken, 1996; Anastasi y Urbina, 1998; Kaplan y Saccuzzo, 2006).

7.5.5. GARANTÍAS PSICOMÉTRICAS Y FUTURO DE LAS TÉCNICAS PROYECTIVAS

Las técnicas proyectivas pese a recibir importantes críticas (falta de sistemas objetivos de valoración, escasos datos sobre fiabilidad y validez, dificultades de puntuación, ausencia de datos normativos...), han seguido utilizándose tanto en el ámbito clínico como escolar (Fernández-Ballesteros, 1991c; Frank de Verthelyi, 1989; Marqués et al., 2004; Meyer y Archer, 2001) desarrollándose, desde algunas hipótesis mencionadas previamente, reactualizaciones de las mismas desde distintos paradigmas teóricos. De las distintas hipótesis planteadas, tres son las vías de desarrollo más importantes y con mayor proyección de futuro:

1. **Reactualización de las técnicas proyectivas desde la perspectiva conductual-cognitiva:** Desde esta perspectiva se pretende un análisis de las variables internas cognitivas con vistas a su análisis funcional y posterior modificación. Con este planteamiento han surgido una serie de técnicas con el fin de evaluar las habilidades cognitivas en la resolución de problemas interpersonales (por ejemplo, el ICPS, Interpersonal Cognitive Problem Solving, véase Spivack y Shure, 1974). Lo importante de este enfoque es que la ejecución de los sujetos en esta variedad de técnicas proyectivas se toma tan sólo como una muestra de la conducta cognitiva de resolver problemas utilizándose, en el caso de déficits, como variable dependiente (junto con otras medidas de ajuste conductual) de un entrenamiento de este tipo de variables. Se ciñen a la evaluación de lo cognitivo, contrastando esto experimentalmente. No obstante, esta forma de proceder elimina uno de los supuestos básicos de estas técnicas, aquel que plantea la importancia de la ambigüedad de los estímulos como base para la proyección. Como señalan Fernández-Ballesteros et al. (1992) el conservar el término proyectivo en este grupo de elaboraciones se debe más bien a razones nostálgicas o de rutina que a planteamientos conceptuales semánticos.
2. **Reactualización de las técnicas proyectivas desde la psicología de la percepción:** Una segunda vía de desarrollos sigue estando en alguna de las corrientes actuales de la psicología de la percepción y, concretamente, en la teoría del procesamiento de la información, la teoría de la decisión, y la teoría de la detección de señales. Desde estas perspectivas se está intentando, a través de la manipulación de las dimensiones estímulares, lograr el esclarecimiento de lo que ocurre desde la presentación del estímulo hasta la formulación de la respuesta. Dos han sido las líneas de investigación más importantes con este enfoque. En primer lugar, la manipulación experimental de los estímulos, del tiempo de exposición y de la consigna con el fin de analizar de modo adecuado las distintas dimensiones y parámetros de las respuestas de los sujetos en base a tales manipulaciones. En segundo lugar, parece necesaria la verificación experimental del tipo de operación que el sujeto realiza internamente a la hora de seleccionar su respuesta ante un estímulo proyectivo. Así, parece necesaria la manipulación experimental de distintos supuestos teóricos (puramente perceptivos, de resolución de problemas, de toma de decisión, etc. (Fernández-Ballesteros et al., 1992).

3. Reactualización de las técnicas proyectivas desde la psicometría: Una última perspectiva futura es la psicométrica, a través de la cual se seguirá intentando la validación de las técnicas proyectivas mediante criterios variados. El trabajo de Exner (1974-1991) sobre el Rorschach es una buena prueba de la posibilidad de contrastación científica de las técnicas proyectivas. Otros estudios analizan su validez concluyendo que el Rorschach es un instrumento válido para evaluar la impulsividad (Pantle, Ebner y Hynan, 1994). En mi opinión, las técnicas proyectivas que quieran sobrevivir deberán estandarizarse según los criterios sugeridos por la psicometría tradicional. En esta dirección se pueden destacar los trabajos (Aguirre, Alamar y Pérez, 1995) que utilizan métodos de corrección objetivos del CAT, tratando de evitar altos niveles de inferencia.

En las dos últimas décadas un grupo de autores han desarrollado un amplio movimiento de reconversión psicométrica de las técnicas proyectivas, intentando no alterar los aspectos sustantivos de las mismas. Desde este enfoque se propone:

- El estudio del valor estimular de los materiales-estímulo proyectivos: láminas, instrucciones.
- La estandarización de las condiciones de administración.
- La obtención de datos normativos.
- La utilización de sistemas de valoración y análisis de las respuestas que sometan a un suficiente control la subjetividad de los evaluadores.
- La interpretación de los resultados dentro de un contexto de contrastación con evidencias procedentes de otro tipo de aproximaciones al estudio de la conducta.

En este contexto de actualización psicométrica se encuentra el T2F, Test de las 2 figuras humanas (Maganto y Garaigordobil, 2009). El objetivo de este trabajo fue crear un instrumento de evaluación madurativo-mental y emocional de niños y niñas de 5 a 12 años con garantías psicométricas. La evaluación madurativo-mental proporciona tablas/criterios de normalidad en el sentido de conocer cuál es el nivel alcanzado por el niño o niña, sujeto de evaluación, entendiendo éste como la posibilidad de detectar a los niños o niñas con sospecha de retraso madurativo, más que en el sentido de diagnosticar el nivel intelectual de éstos. La evaluación emocional proporciona datos que permiten conocer las principales preocupaciones de los niños y niñas en estas edades, sus temores, miedos, ansiedades y otros aspectos de la experiencia emocional del sujeto que inciden favorable o desfavorablemente en su comportamiento. Es un test de screening en una primera aproximación diagnóstica, para detectar sujetos susceptibles de una evaluación más específica. Tiene también una utilidad clínica en el sentido de alertar sobre la necesidad de proceder a una evaluación más completa en el caso de presentar importantes problemas madurativos y/o emocionales. La tarea solicita hacer dos dibujos de la figura humana, una masculina y otra femenina. La evaluación madurativo-mental se basa en la asignación de puntuaciones estandarizadas a cada uno de los ítems que componen la evaluación de las dos figuras humanas, en concreto 52 indicadores madurativos (IM). Y la evaluación emocional en la identificación de indicadores emocionales (IE) de una lista de 35 indicadores definidos. Los estudios psicométricos de validez y fiabilidad han confirmado las garantías psicométricas de esta prueba proyectiva gráfica.

Otros trabajos han aplicado los conocimientos derivados de la TRI a la investigación de algunas técnicas proyectivas. Por ejemplo el estudio de Fermíno Fernandes (2007) tuvo como

objetivo la determinación del funcionamiento diferencial del ítem (DIF) en el test del Dibujo de la Figura Humana utilizado para la evaluación de la inteligencia en niños, considerando la variable sexo. Se estudiaron los dibujos de 2508 estudiantes de la enseñanza primaria y preescolar, cuyo promedio de edad fue de 8,14 años. Gran parte de ellos frecuentaban escuelas públicas (72,1%) y los grupos para el estudio de DIF fueron compuestos por 1248 niños y 1260 niñas. Los parámetros métricos de los ítems fueron determinados según el modelo logístico de un parámetro de Rasch. Los resultados indicaron la presencia de DIF en 20 de los 46 ítems analizados, 9 de ellos han sido de fácil ejecución para las niñas y 11 para los niños.

Finalmente cabe recordar las sugerencias sobre fiabilidad y validez de las técnicas proyectivas que se han realizado en los estándares de la APA (Avila, 1997; Valdés y Gómez de Terreros, 2003).

1. No tiene ningún sentido pretender validar una prueba proyectiva como un “todo”, sino que es más adecuado especificar y referirse a ella en función de los objetivos planteados, procedimiento seguido... e indicar si se ha mostrado válida para predecir determinadas variables comportamentales.
2. Cada instrumento proyectivo posee características y problemas propios. Esto significa que las evidencias (positivas o negativas) de fiabilidad y validez no pueden ser extrapoladas de unos instrumentos a otros, como tampoco mucho más allá de las condiciones experimentales en las que fueron obtenidas.
3. Es necesario tener en cuenta, sobre todo a la hora de emplear estas técnicas para la investigación, la modulación que introducen variables intervinientes tales como edad, sexo, maduración fisiológica, nivel educativo y socioeconómico así como las interacciones de todos estos factores.
4. Cuando es necesario el diagnóstico diferencial, se debe escoger un test para el cual haya evidencia de su aptitud para distinguir entre dos o más grupos de diagnósticos, más que para distinguir los casos patológicos de la población general.
5. El problema de la validez ha de abordarse mediante múltiples métodos combinados.
6. No deben utilizarse respuestas específicas de un individuo ante un estímulo proyectivo como base suficiente para la evaluación psicológica, a menos que conste evidencia de la validez de tal inferencia.
7. No debe contraponerse ni considerarse excluyentes los usos clínicos y actuariales de las técnicas proyectivas. Una actuación combinada de procedimientos clínicos y actuariales puede optimizar su praxis. La validez y fiabilidad de estas técnicas depende de un uso contextual adecuado en el que se utilicen combinadas toda clase de evidencias.

La solidez psicométrica de muchos instrumentos proyectivos usados en forma amplia todavía está por demostrarse. El debate entre académicos que afirman que las pruebas proyectivas no son instrumentos sólidos desde el punto de vista técnico y los psicólogos clínicos que encuentran útiles estas pruebas ha sido incontenible incluso desde que las pruebas proyectivas comenzaron a ser usadas ampliamente (Cohen y Swerdlik, 2001). Recientes estudios han evidenciado que las técnicas proyectivas, como el Rorschach y TAT siguen estando en uso y, para muchos profesionales, son instrumentos valiosos en el diagnóstico y evaluación de los trastornos de la personalidad (Rodríguez Sutil, 2007). No obstante, para que su uso sea adecuado deberán estar integradas dentro del más complejo proceso diagnóstico que es donde deben ser valoradas.

7.6. TÉCNICAS SUBJETIVAS

7.6.1. DEFINICIÓN Y CARACTERÍSTICAS DE LAS TÉCNICAS SUBJETIVAS

Al hablar de técnicas subjetivas en evaluación psicológica hacemos referencia a una serie de métodos orientados a recoger y organizar información acerca de la estructura psicológica, los contenidos y los procesos de las visiones subjetivas o significados personales acerca de uno mismo y el mundo (Feixas, 2003). En contraste con los métodos que tienen como objeto clasificar al sujeto en unas determinadas dimensiones (por ejemplo extraversión-introversión) o categorías (trastorno mental) ya establecidas de antemano, las técnicas subjetivas pretenden captar sus construcciones personales, en definitiva, estudiar sistemáticamente la estructura de su subjetividad (Feixas y Botella, 2003).

Tradicionalmente, las técnicas subjetivas han surgido en el seno de teorías u orientaciones psicológicas con cierto cariz fenomenológico, que defienden el valor de las construcciones subjetivas del sujeto a la hora de comprender su conducta y emociones. Los enfoques constructivistas (Feixas y Villegas, 2000) serían un ejemplo paradigmático, que en el campo de la evaluación psicológica ha generado su propia corriente (Neimeyer, 1993).

Las técnicas de calificación, también llamadas subjetivas, comprenden aquellos procedimientos de recogida de información que presentan las siguientes características (Fernández-Ballesteros, 1992e):

- La fuente de datos puede ser el propio sujeto que se califica o clasifica a sí mismo o a objetos, personas o conceptos significativos, o también personas allegadas al sujeto en exploración que realizan esa misma tarea pero referida a éste.
- El tipo de material suele ser semiestructurado y no enmascarado. El material no está tipificado, es flexible y se adapta a los objetivos que se persiguen en cada caso.
- La respuesta del sujeto es voluntaria y, por lo tanto, puede ser falseada.
- El tipo de tarea requerida es la de descripción, calificación o clasificación del propio sujeto o de otros objetos o abstracciones.
- La manipulación de los resultados no está reglada, pudiendo optar por planteamientos cualitativos o cuantitativos en la elaboración de respuestas.
- Estas técnicas han sido construidas desde planteamientos teóricos fenomenológicos y/o cognitivos, aunque en la actualidad se emplean desde perspectivas teóricas muy distintas (conductismo, psicoanálisis...).

Así, estas técnicas pueden ser consideradas autoinformes, ya que son técnicas de autocalificación y autclasificación. Avila (1990) que incluye los autoinformes en las técnicas subjetivas señala tres características comunes a las mismas:

- Utilizar como estímulos etiquetas verbales descriptores de la emoción, motivación, personalidad u otros estados internos (contenidos cognitivos).
- Utilizar respuestas personales, subjetivas, bien por elección-autoatribución de dichas etiquetas verbales, o bien por efectuar juicios de estimación sobre las mismas.
- Sometimiento de las respuestas emitidas a un tratamiento psicométrico que permita establecer indicadores cuantitativos y efectuar comparaciones intra e intersujetos.

Como se puede observar, ambos autores discrepan en esta última característica ya que para Fernández-Ballesteros (1992e) la manipulación de los resultados no está reglada, pudiendo optar por planteamientos cualitativos o cuantitativos en la elaboración de respuestas.

7.6.2. TÉCNICAS SUBJETIVAS REPRESENTATIVAS

A la hora de intentar una clasificación de las técnicas subjetivas debemos constatar que existe una dificultad en cuanto al uso del término "subjetivas". Así, la entrevista, uno de los procedimientos de evaluación más importante, sería para algunos (Pervin, 1975/1979) una técnica subjetiva, mientras que otros la incluyen dentro de los autoinformes aunque la consideran con matices diferenciales respecto a estos (Fernández-Ballesteros, 1980, 1992e). Pese a estas discrepancias, generalmente se consideran más representativas las siguientes:

- Técnica de Clasificación Q.
- Diferencial Semántico.
- Test de los Constructos Personales.
- Listas de Adjetivos.
- Método Biográfico.

7.6.2.1. La técnica de clasificación o de distribución "Q" de Stephenson (1953)

Consiste en un procedimiento de evaluación basado en organizar tipologías de rasgos de personalidad bien para describir sujetos o grupos de sujetos, y aplicable al propio sujeto o a personas allegadas a él. La técnica de clasificación "Q" ha sido utilizada para el estudio del autoconcepto (Fierro, 1982) y en la descripción de la personalidad de un sujeto con fines de análisis. Dotada de alta precisión, está destinada fundamentalmente a la investigación. Sus características más relevantes son (Fernández-Ballesteros, 1992e):

- El material no está previamente establecido, depende de los objetivos del investigador y está constituido por una muestra del universo de elementos sobre los que se trate de efectuar el análisis.
- El universo de tales elementos sobre los que se selecciona la muestra es, generalmente un conjunto de manifestaciones o descripciones expresadas en forma verbal-escrita. No obstante, pueden construirse clasificaciones-Q sobre objetos, abstracciones...
- La tarea que se pide al sujeto es que autoaplique tales manifestaciones según la frecuencia de aparición o su aplicabilidad, bien a sí mismo, bien a otro sujeto u objeto conocido del cual se pretende obtener información.
- Las alternativas de puntuación son de elección forzosa y, por tanto, deben ajustarse a la curva normal.
- Las condiciones de aplicación dependen de los objetivos del investigador.
- El tratamiento estadístico de la clasificación-Q está en dependencia de la finalidad con la que se hayan recogido los datos, aunque debería ajustarse a un diseño intraindividual.

Han aparecido críticas a la metodología Q propuesta por Stephenson (Edwards, 1955), aunque la técnica ha tenido derivaciones notables. Una de ellas, es la revisión de la Clasificación "Q" conocida como "Clasificación-CQ" o "California Q-set" que consiste "en un

conjunto de descripciones de personalidad que han de ser ordenadas en forma tal que describan lo más exactamente posible, a la persona objeto de estudio” (Block, 1978), con el objeto de que sea posible la comparación cualitativa y el análisis de la misma (Rodríguez Franco y Blanco Picabia, 1986). Fue desarrollada por Block (1961-1978) y, ha sido adaptada al español en la Universidad Autónoma de Madrid (Fernández-Ballesteros, 1992e).

7.6.2.2. La Técnica del Diferencial Semántico de Osgood (1952)

Se trata de un instrumento de evaluación que aparece en el contexto de la teoría mediacional de Osgood. Según esta teoría las palabras representan cosas porque producen en los organismos humanos algún tipo de réplica de su conducta con respecto a esas cosas como proceso mediacional. El interés de Osgood fue estudiar la estructura del significado, para lo cual realizaron una serie de experimentos en los que se fundamenta el diferencial semántico.

La técnica del Diferencial Semántico ideada por Osgood (Osgood, 1952; Snider y Osgood, 1972) pretende estudiar la significación que personas, eventos, conceptos... tienen para determinado individuo (Rodríguez Franco y Blanco Picabia, 1986). Esta técnica mide las respuestas de los sujetos frente a objetos o estímulos semánticos llamados conceptos, a través de escalas de estimación definidas por adjetivos bipolares según un formato previamente definido. Así, el DS requiere: una serie de estímulos semánticos o conceptos sobre los que se desea indagar su significado en un sujeto o grupo de sujetos, unas escalas de estimación definidas por adjetivos bipolares, y la elección de un formato a utilizar. Esta técnica está basada en la llamada Teoría Mediacional de Osgood, según la cual existen procesos intermedios encubiertos que regulan la relación funcional entre estímulo y respuesta.

Como señala el propio Osgood, el Diferencial Semántico no es un instrumento de medida que tenga un número concreto de ítems, ni tampoco una puntuación específica, sino que muy al contrario, es una técnica generalizada de medida que posee gran flexibilidad y que debe ser adaptada en función de los propósitos de la investigación que se quiera utilizar (Atienza y Fuentes, 1994). Así los conceptos y escalas que se utilicen serán seleccionados por el evaluador. Habitualmente para la construcción de un DS se dan los siguientes pasos: selección de conceptos, selección de escalas, selección del formato, administración, puntuación y análisis de datos.

Se han elaborado escalas correspondientes al DS en la investigación del concepto del “Yo” y del “sí mismo” (Fierro, 1982). Existen otras variantes como el *Inventario Diferencial de Adjetivos para el Estudio del Estado de Animo (IDDA-EA)* de Tous y Andrés (1991), Diferencial Semántico construido con el fin de medir el estado de ánimo teniendo en cuenta las tres dimensiones o factores (activación, estrés y arousal) de la autopercepción individual.

7.6.2.3. El Test de los Constructos Personales (REP) de Kelly (1955)

El marco teórico de la prueba es la teoría de los constructos personales de Kelly (1955). El test de los constructos personales es una técnica de evaluación subjetiva que se desarrolló con el objetivo de evaluar los constructos que las personas utilizan para interpretar su entorno (Atienza y Fuentes, 1994), derivándose directamente de su teoría de la personalidad. Al igual que consideraba Osgood, Kelly mantiene que el significado de las cosas no es inherente al objeto en sí mismo, sino que emana de la persona. De esta manera, el individuo se implica

activamente en la interpretación de los acontecimientos (Rodríguez Franco y Blanco Picabia, 1986).

Este test fue diseñado por Kelly con el objetivo de dar cuenta de cómo las personas construyen los distintos roles interpersonales con el fin de establecer hipótesis clínicas sobre la persona objeto de estudio. El REP permite el estudio del lenguaje del sujeto referido a su mundo, preferentemente interpersonal. Los desarrollos posteriores a los que ha dado lugar el REP ofrecen, por su flexibilidad y posibilidad de tratamiento matemático de los resultados, una serie de ventajas que permiten su utilización no sólo en el planteamiento de hipótesis terapéuticas sino también en la investigación básica y aplicada.

En su modalidad interpersonal (la más utilizada), la persona entrevistada identifica a personas significativas de su entorno (elementos de la rejilla). A continuación se buscan similitudes y diferencias entre pares de estos elementos (constructos). Obtenidos los elementos y constructos, la persona puntúa cada elemento en función de cada constructo, mediante una escala Likert (usualmente de 7 puntos) (Feixas y Botella, 2003).

Este instrumento puede ser considerado como uno de los más vinculados a una teoría. Sin embargo, se utiliza en múltiples ocasiones sin relacionarlo con la teoría de los constructos personales de la cual surgió. La no estructuración del instrumento ha llevado a bastantes autores a defender la hipótesis de que los criterios de fiabilidad y validez utilizados normalmente para valorar un instrumento plantean ciertas dificultades al intentar ser aplicados a esta técnica subjetiva. No obstante, Rivas y Marco (1985) después de realizar una revisión de varios trabajos concluyen: (1) La validez de la rejilla para mostrar las relaciones entre constructos; (2) La sensibilidad de la rejilla para reflejar cambios; y (3) La validez de la rejilla para hacer predicciones.

En los últimos años, se ha operativizado una aplicación informatizada de este instrumento que permite evaluar las dimensiones y estructura psicológica del constructo “significado personal”, aspectos de la forma en que una persona da sentido a su experiencia y organiza el mundo (Díez, Gironella y Feixas, 1993; Feixas y Cornejo, 1992, 1996; Feixas, Cornejo y Bach, 1991; Feixas y Botella, 2003).

7.6.2.4. Las Listas de Adjetivos (LA)

Las LA consisten en listados de adjetivos que se presentan a un sujeto con el fin de que juzgue si pueden ser aplicables, o bien a sí mismo, o bien a una persona u objeto del cual se le pide juicio. La apreciación del sujeto suele ser recogida en forma numérica a través de escalas de puntuación que a su vez pueden presentar diferentes formas (escalas gráficas, de elección forzosa, acumulativas, de intensidad y de orden) (Atienza y Fuentes, 1994; Fernández Ballesteros, 1992e; Rodríguez Franco y Blanco Picabia 1986).

Las listas de adjetivos pueden estar referidas a muy distintos campos de aplicación. Avila y Giménez de la Peña (1991) han revisado las distintas posibilidades de aplicación, así como los diferentes modelos en los que pueden ser encuadrados los procesos emocionales cognitivos y conductuales que requieren. Se recomienda al lector interesado en esta forma de calificación consultar este trabajo.

Estas LA pueden o no estar estandarizadas, algunas son construidas para cada caso, mientras que otras disponen de datos normativos. De ellas, las más representativas son: La Lista de Adjetivos (ACL) de Gough y Heilbrum (1965-1980), y La Lista de Adjetivos de

Depresión (DACL) de Lubin (1967-1981). Ambas arrojan buenos índices de fiabilidad y validez, siendo ampliamente utilizadas en investigación. En España Avila Espada (1992c) está realizando un importante esfuerzo de adaptación y diseño de listas de adjetivos para múltiples propósitos (por ejemplo, de la DACL o de la IAS-R), operándose diversos análisis de las mismas (Avila Espada, 1990b; Avila Espada y Giménez de la Peña, 1991ab; Avila Espada, Felipe y Gozalo, 1994).

7.6.2.5. El método Biográfico

Es especialmente importante cuando el objetivo de la evaluación es el análisis extensivo e intensivo de la personalidad individual. Las técnicas más habituales son los “*Life Events*” y los “*Inventarios Biográficos*”. El examen sistemático de los acontecimientos o sucesos de la vida diaria es relativamente un nuevo desarrollo, aunque hay una larga y fragmentada historia del estudio de estos sucesos o experiencias, y se contraponen al estudio tradicional de los resultados vitales (Ozer y Reise, 1994). En la investigación sobre las experiencias de la vida se observan tres líneas de interés: (1) la vinculación de los sucesos de la vida diaria con afectos y sentimientos de bienestar; (2) la relación de estos eventos a explícitas metas, motivos e intenciones; y (3) estudio de las relaciones entre experiencias diarias de la vida y rasgos de personalidad.

7.7. LAS TÉCNICAS OBJETIVAS

Desde Galton, Binet y McKeen Cattell, la evaluación de las respuestas senso-perceptivas y motoras mediante instrumentación psicológica ha sido de gran importancia para la psicología; siguiendo esa tradición, actualmente las técnicas objetivas evalúan especialmente este tipo de respuestas. Por ello, se aplican a aspectos perceptivos y sensoriales que pueden o no implicar para su evaluación respuestas motoras. Durante mucho tiempo la tecnología utilizada se ha basado en aparatos mecánicos que, procedentes de los primeros laboratorios de psicología, en la práctica apenas se utilizaban. Los avances de la tecnología han permitido un extraordinario desarrollo desde esa instrumentación mecánica hasta un conjunto de técnicas electrónicas y/o informatizadas que se agrupan con la rúbrica de técnicas objetivas (Calero y Fernández-Ballesteros, 2004).

Actualmente, estas técnicas han ampliado su uso a nuevas áreas de aplicación y hoy día no sólo son utilizadas en el laboratorio para el estudio de procesos cognitivos básicos, sino que, también, se consideran muy útiles en la evaluación de determinados problemas neuropsicológicos, en la valoración de los efectos de la ingestión de determinados medicamentos o drogas, en la identificación de algunos síntomas clínicos (ansiedad, demencia...) y, por último, la evaluación necesaria para la obtención y revisión del permiso de conducción, de armas, en la selección profesional especializada, en deporte de élite, y en otros muchos ámbitos aplicados...

7.7.1. DEFINICIÓN Y CARACTERÍSTICAS DE LAS TÉCNICAS OBJETIVAS

La categoría de técnicas objetivas ha sido utilizada para encuadrar muy variados instrumentos, a veces casi incompatibles por sus características. Catell (1965) considera como tests objetivos diferentes tipos de pruebas (proyectivas, de ejecución, de expresión gráfica y

psicofisiológicas), planteando que pueden definirse, según el criterio de que el sujeto no sabe en qué aspecto de su conducta se le está evaluando. Eysenck (1959) considera como tests objetivos aquellas tareas estructuradas que permiten recoger la conducta motora externa de los sujetos y cuya puntuación es independiente del evaluador, es decir, incluye los tests de destreza motora, velocidad de golpeo, pruebas de trazado, etc. Para Pervin (1975/1979) en la categoría objetiva se incluyen aquellas pruebas que tienen la cualidad de ser estructuradas, contar con un número limitado de alternativas de presentación y llevar consigo que la respuesta no puede ser modificada voluntariamente. Así, incluye las pruebas de ejecución que se realizan en situaciones estandarizadas de laboratorio, la observación directa, y los registros psicofisiológicos. Dieguez Sánchez y Blanco Picabia (1986) consideran técnicas objetivas los cuestionarios, inventarios y escalas observacionales, así como los contruidos por una serie de aparatos en su mayoría de estructura mecánica. Entre los cuestionarios se descartan aquellos que implican autoobservación del sujeto y, únicamente, se aceptan los sistemas de codificación utilizados en observaciones (naturales y artificiales) cuya estructura de formato limitan al máximo la capacidad de acción del evaluador.

Como ya he señalado previamente, las clasificaciones de las técnicas son variables y deben ser exhaustivas y excluyentes, sin embargo los distintos autores utilizan diferentes parámetros para su clasificación lo que provoca que un instrumento se encuentre agrupado de forma diferente en las categorizaciones realizadas por distintos investigadores.

Siguiendo la estructuración planteada por otros autores (Calero y Fernández-Ballesteros, 2004; Fernández-Ballesteros y Calero, 1992), en este proyecto definimos las técnicas objetivas como: aquellos procedimientos de recogida de información de eventos psicológicos observables o amplificables, que en gran parte de los casos no son controlables (al menos sin entrenamiento), los cuales se aplican mediante sofisticados aparatos que permiten una administración, registro, puntuación y análisis objetivo (generalmente mecánico o electrónico) sin la intervención del evaluado". Tres son pues sus características principales:

- Requieren una instrumentación y un material estándar cuya aplicación se realiza en condiciones estructuradas y de máximo control (objetividad en el procedimiento y artificialidad).
- El sujeto no puede modificar sus respuestas de forma voluntaria.
- Las respuestas del sujeto pueden ser registradas, codificadas y procesadas sin que medie la opinión o criterio del evaluador.

Además de estas características específicas las técnicas objetivas suponen una variedad de la observación (observación en situaciones artificiales), se sitúan básicamente dentro de una metodología experimental y se basan en los principios de la Psicología del Aprendizaje y la Psicología Cognitiva.

7.7.2. TIPOS DE TÉCNICAS OBJETIVAS

Siguiendo la primera clasificación realizada por Fernández-Ballesteros y Calero (1992) las técnicas objetivas se clasifican tomando como criterio la variable que el instrumento evalúa, destacando las técnicas objetivas en los dominios cognitivo, motor y psicofisiológico. Posteriormente, estas investigadoras incluyen de forma diferenciada las técnicas objetivas perceptivas (Calero y Fernández-Ballesteros, 2004). Además, tanto en la categorización que

realizaron en 1992 como en la llevada a cabo más recientemente (Calero y Fernández-Ballesteros, 2004) incluyen en esta categoría los instrumentos aplicados por ordenador debido a que la computadora permite objetivar la administración, corrección y puntuación y, en algunos casos la interpretación. No obstante, las autoras subrayan el hecho de no confundir estos tests aplicados por ordenador con las técnicas objetivas manifiestas o amplificadas sobre las que el sujeto tiene una relativa falta de control. Tomando como criterio la variable evaluada por el instrumento, las técnicas objetivas se pueden clasificar en cuatro categorías: perceptivas, cognitivas, motoras, y psicofisiológicas.

7.7.2.1. Instrumentación Perceptiva

Dentro de esta categoría se incluyen aparatos que permiten evaluar la percepción táctil, auditiva, visual... (Calero y Fernández-Ballesteros, 2004).

En relación con la *percepción táctil*, la variable de mayor interés en evaluación psicológica en este momento es la sensibilidad, relacionada o no con el dolor, que es importante tanto en evaluación neuropsicológica como en valoración de algunas intervenciones rehabilitadoras en dolor crónico y/o recuperación funcional. Algunos ejemplos de estos aparatos son:

- *Estesiómetro*: instrumento que mediante una fibra de vidrio y un sensor registra la sensibilidad de la piel en un punto. Se utiliza para la evaluación de problemas neurológicos.
- *Test de discriminación de dos puntos*: igual que el anterior sólo que éste establece la comparación entre dos puntos diferentes a la vez. Se utiliza para evaluar la sensibilidad táctil.
- *Dolorímetro electrónico portátil*: instrumento que registra el umbral del dolor mediante presión en la superficie de la piel a través de un sensor. Se utiliza cuando se existen problemas de sensibilidad al dolor.

En relación con la *percepción auditiva*, existen una serie de instrumentos para evaluar la discriminación auditiva que, en general, se dirigen a distinguir la discriminación de sonidos ambientales, discriminación figura-fondo y discriminación de palabras. Suelen acompañarse de la presentación de imágenes, mediante ordenador o en formato libro, en las que el sujeto evaluado debe señalar el origen del sonido. Estos instrumentos poseen una máxima utilidad en trastornos auditivos, del habla y del lenguaje y en rehabilitación logopédica; también se utilizan en relación con la determinación de la afasia, la inatención auditiva y/o la escucha dicótica. Algunos ejemplos de estos aparatos son:

- *Sonómetro*: medida automática y directa de la precisión auditiva en decibelios. Se utiliza para la evaluación de problemas auditivos y/o de lenguaje.
- *Test de reconocimiento de sonidos*: test formado por una grabación de sonidos y la presentación de imágenes, informatizadas o en papel, para evaluar de diferentes modos la discriminación auditiva. Permite la manipulación de diferentes parámetros: familiaridad de los sonidos, discriminación figura/fondo, intensidad, etc.

Con respecto a la *percepción visual*, las variables más frecuentemente examinadas mediante técnicas objetivas son: el análisis visual de estímulos, los modelos de análisis visoperceptivos y los juicios sobre aspectos espaciales del estímulo. Para ello existen actualmente diferentes tipos de aparatos dirigidos a la evaluación de distintos aspectos tales como: vigilancia, reconocimiento visual, discriminación, memoria, relaciones espaciales, constancia de formas, memoria secuencial, figura-fondo y cierre, visión en tres dimensiones, punto de fusión, discriminación luminosa, percepción de la profundidad, percepción horizontal-vertical, velocidad de anticipación, percepción del movimiento, etc. Todos ellos resultan muy interesantes en neuropsicología para la valoración de lesiones o problemas cerebrales y cuando se trata de comprender las diferencias funcionales del cerebro. También se utilizan en evaluación de los efectos de determinadas drogas o patologías sobre las funciones neuropsicológicas y en evaluación para permiso de conducir, de armas o en selección profesional específica. Algunos ejemplos de estos aparatos son:

- *Medida de la Percepción de la Profundidad:* Se evalúa con un tipo de aparato que consta de un cajón con un visor, en cuyo interior iluminado se presentan al sujeto dos varillas verticales, una fija y otra móvil que puede ser accionada a distancia. El sujeto percibe ambas varillas a través del visor y su tarea consiste en colocar la varilla móvil a la misma distancia de él que la fija. Se emplea comúnmente en la selección de conductores.
- *Percepción de la verticalidad. El test del Marco y la Varilla:* Conocido por el Rod-and-Frame Test (RFT) es uno de los instrumentos más utilizados en la percepción de la verticalidad. Consta de un marco y una varilla luminosos y móviles conectados a un dispositivo en el que se registran los grados de desviación de la varilla sobre la verdadera vertical, el cual permite accionar a distintos grados tanto el marco como la varilla. La tarea del sujeto consiste en situar verticalmente la varilla, que es presentada con diferentes grados de inclinación dentro del marco igualmente inclinado. En lo que se refiere al ámbito de la personalidad este instrumento ha sido utilizado en la evaluación del estilo cognitivo “dependencia-independencia de campo”.
- *Aparato ilusiones ópticas:* Aparato que presenta diversos efectos visuales, tales como ilusiones de profundidad, ángulo, movimiento y color, para medir la precisión visual de estos fenómenos.
- *Aparatos para la discriminación luminosa:* Aparato que presenta diferentes estímulos luminosos para que el sujeto estime cuál es el más intenso.
- *Medidor de velocidad de anticipación y percepción del movimiento:* El sujeto observa una luz desplazándose y debe estimar en qué momento llega al final establecido. Se dirige a evaluar la percepción de velocidad.

7.7.2.2. Instrumentación Cognitiva

Con la denominación instrumentación cognitiva se hace referencia a todos aquellos aparatos que se dirigen a evaluar variables cognitivas consideradas observables (aunque fuera mediante amplificación de una señal); específicamente, a medidas de atención, memoria y procesamiento de la información. La mayoría de estas variables pueden ser recogidas de manera precisa con la utilización de diversos aparatos y también con diversos programas informáticos diseñados para tal fin. Ambos formatos permiten la manipulación de las condiciones de presentación, de los estímulos, tiempos de presentación y niveles de complejidad de la tarea, suministrando, asimismo, el registro automático de las respuestas y de los errores según diversos parámetros: aciertos, errores, tiempos de ejecución, latencias, etc. Así, existen diversos aparatos dirigidos a evaluar la atención entendida como capacidad limitada relacionada con el esfuerzo sostenido de focalización según 4 componentes básicos: selección sensorial, selección de respuesta, capacidad atencional y ejecución sostenida (Larrabee, 2000).

Todos estos instrumentos incluidos en esta categoría permiten la manipulación de las condiciones físicas estimulares en la presentación, y registro automático de las respuestas de los sujetos en base a ciertas dimensiones de las respuestas como los aciertos, errores, así como según parámetros temporales tales como el tiempo de latencia, tiempo de ejecución, tiempo de error, etc. (Calero y Fernández-Ballesteros, 2004; Fernández-Ballesteros y Calero, 1992). Algunas de las técnicas más representativas son:

- *Polirreactígrafo*: Permite la medición automática de todas las modalidades de tiempos de reacción, ya que tiene programados tiempos de reacción simples, auditivos y visuales, las reacciones de elección auditivas y visuales, la estimulación difusa, la atención concentrada, y las reacciones complejas. El aparato presenta estímulos luminosos y auditivos al sujeto, a los que debe responder con los dispositivos de respuestas que son dos manillas una para cada mano y dos pedales para los pies. Se emplea en la investigación básica sobre la atención, en la selección de profesionales que requieran habilidades de atención y vigilancia (controladores aéreos...), en psicología deportiva (evaluación de deportistas de élite) y en la evaluación clínica (déficits cognitivos atencionales de los esquizofrénicos...).
- *Taquistoscopio*: Este aparato dispone de un presentador de estímulos a los que el sujeto debe de dar respuesta cuando aparece el estímulo convenido. Es el aparato más importante en la investigación de la actividad cognitiva, y específicamente en la investigación sobre percepción y memoria. En el campo de la Evaluación psicológica se ha empleado para investigar el procesamiento icónico de los estímulos del Rorschach con el fin de comprobar determinadas hipótesis sobre la presencia de los determinantes de las respuestas a las láminas del Rorschach (Márquez y Fernández-Ballesteros, 1988). También se ha utilizado para la evaluación del razonamiento analógico siguiendo la aproximación de análisis de componentes (Calero, 1984, 1989). Otro núcleo de estudios procede de la investigación del déficit cognitivo esquizofrénico, la cual se ha realizado en su mayor parte mediante taquistoscopio (Castilla del Pino, Ruiz Vargas, Adarraga, Belinchón, y Vizcarro, 1991). Sin embargo, como señala Vizcarro (1991), medidas perceptivas objetivas derivadas de los estudios sobre la esquizofrenia están todavía muy

lejos de tener aplicabilidad. Un último ámbito de aplicación de este aparato es el de la evaluación psiconeurológica (Corbera y Grau, 1993).

- *Tambor de memoria*: Rotor que contiene diferentes estímulos que se presentan al sujeto por un breve espacio de tiempo y él debe recordar. Permite trabajar con diferentes estímulos, que se combinan aleatoriamente, y fijar el tiempo de presentación.

7.7.2.3. Instrumentación Motora

Una de las funciones más estudiadas en evaluación psicológica es la función motriz y, especialmente la denominada coordinación psicomotriz. Esta función hace referencia a la precisión, rapidez y organización espacial puestas en práctica en la ejecución de determinadas tareas. Estos aparatos permiten la administración, puntuación e incluso a veces, procesamiento de ejecuciones psicomotoras ante tareas de laboratorio en contextos clínicos y deportivos (Calero y Fernández-Ballesteros, 2004; Fernández-Ballesteros y Calero, 1992). Las respuestas motoras más evaluadas son: funciones táctiles, coordinación manual, bimanual y óculo-manual, estabilidad y tiempo de anticipación. Su aplicación suele darse en relación con la valoración de lesiones cerebrales, problemas cerebro-vasculares y/o tumores que pueden afectar a los sentidos: audición, visión, tacto...; también en la valoración de algunas enfermedades como el Parkinson y en la valoración de tratamientos tales como dolor crónico músculo-esquelético o temblores. Otras áreas importantes de aplicación son la psicología deportiva y el reconocimiento psicotécnico de conductores. Algunos de los más empleados en el campo de la evaluación son:

- *Omega*: Mide la capacidad para disociar los movimientos de ambas manos, estando compuesto por una placa metálica con una ranura en forma de omega, en cuyo interior se mueve un pivote metálico accionado mediante dos mandos cuyos ejes forman un ángulo con la placa variable centrado en 30°. La tarea del sujeto consiste en desplazar el pivote de un extremo a otro del recorrido, sin tocar los bordes de las ranuras. En el campo de la evaluación se ha utilizado especialmente en la selección de oficios mecánicos.
- *Precisión Manual de Purdue*: Es un aparato de medición de la habilidad visomotora. Consta de una plataforma con un orificio en la que un círculo, con orificio coincidente, gira a distintas velocidades, cuenta también con un controlador de velocidad y un punzón. El sujeto ha de mantener el punzón mientras gira a distintas velocidades. Este aparato ha sido utilizado en la selección profesional de oficios que requieren finas habilidades psicomotrices.
- *Octobinexímetro-Tremómetro*: Este instrumento permite la evaluación de un conjunto amplio de habilidades manuales, planteando diversas tareas que permiten explorar distintas funciones psicomotrices como: rapidez, precisión, seguridad, relación rapidez-precisión, capacidad de mejora de la precisión y del rendimiento, así como el índice de ambidextría.

- Otros aparatos de evaluación de variables psicomotrices: *Simulador de conducción (coordinación motriz), silla vibrátil (actividad)...*

7.7.2.4. Instrumentación Psicofisiológica

La utilización de técnicas objetivas en el registro de respuestas psicofisiológicas ha sido utilizada principalmente desde la Psicofisiología, rama de la Psicología centrada en el estudio de las relaciones entre la fisiología del organismo y la conducta. Especialmente han sido técnicas bastante utilizadas desde el enfoque conductual, realizándose varias sistematizaciones sobre las mismas (Calero y Fernández-Ballesteros, 2004; Fernández-Ballesteros y Calero, 1992; Matud, 1993d; Vila, 1994). Generalmente, estos aparatos tienen un proceso de registro de las respuestas psicofisiológicas que pasa por cinco fases: (1) Detección de la señal proveniente del organismo; (2) Transformación de la señal del organismo en señales eléctricas; (3) Amplificación de las mismas; (4) Registro propiamente dicho; y (5) Conversión de la señal registrada en formas que facilitan su análisis y, en ocasiones, propiamente análisis estadístico.

Estos instrumentos tienen por finalidad explorar diversas respuestas psicofisiológicas. Respecto a las principales respuestas psicofisiológicas, existen múltiples sistemas de clasificación. Si se ordenan en función del sistema que las emite, se pueden diferenciar:

- Sistema Somático: respuesta electromiográfica, movimientos oculares, respiración.
- Sistema Nervioso Autónomo: sistema cardiovascular (tasa cardíaca, presión y flujo sanguíneo), temperatura corporal, excitación sexual (pletismografía del pene y la vagina), respuestas electrodérmicas (resistencia, conductancia y potencial de la piel...), respuesta pupilográfica, respuestas gastrointestinales, salivación.
- Sistema Nervioso Central: respuestas electroencefalográficas, respuestas evocadas.
- Sistema Endocrino y Bioquímico.

En similar dirección, Vila (1994) clasifica las técnicas de registro psicofisiológico de acuerdo con la organización estructural y funcional del sistema nervioso, planteando 3 categorías:

- Sistema Nervioso Central: actividad electroencefalográfica, potenciales evocados, cartografía de la actividad eléctrica cerebral, tomografía por emisión de positrones.
- Sistema Nervioso Somático: actividad electromiográfica, movimientos oculares, actividad respiratoria, medidas indirectas de la actividad somática.
- Sistema Nervioso Autónomo: actividad electrodérmica, actividad cardiovascular, actividad pupilar, actividad salivar, actividad gastrointestinal, actividad sexual.

Las principales respuestas psicofisiológicas que han sido estudiadas son: la actividad eléctrica de la piel, la actividad cardiovascular, la actividad electromiográfica y la actividad cerebral (Vila, 1994). Respecto a los aparatos de registro de las respuestas psicofisiológicas, Fernández-Ballesteros y Calero (1992) destacan tres instrumentos empleados con mayor frecuencia:

- *El polígrafo:* Es el más frecuentemente utilizado a la hora de registrar respuestas psicofisiológicas. Es un aparato mediante el cual se detectan, amplifican y reproducen las señales procedentes de los distintos sistemas de respuestas fisiológicas. Dispone de tres partes: (1) la unidad de detección de señal; (2) la unidad de amplificación y/o procesamiento de la señal; y (3) la unidad de transducción, que convierte la señal eléctrica en una inscripción sobre un soporte no volátil, para facilitar el posterior análisis de la respuesta.
- *Sistemas modulares computerizados:* En la actualidad existen aparatos mucho más modernos que permiten el registro, transformación, amplificación y el cómputo de múltiples respuestas psicofisiológicas. Un ejemplo de este tipo de aparatos es el BIOPAC, que es un sistema modular de fiobeedback y control fisiológico, compuesto por módulos electrofisiológicos de doble canal, así como del software necesario para realizar el control necesario mediante un ordenador personal. Estos módulos transfieren la información psicofisiológica a una computadora para su almacenamiento en memoria y posterior utilización para obtener las gráficas de las mismas, su análisis estadístico...
- *Aparatos de Registro Portátiles:* Consisten en pequeños aparatos para ser utilizados en situación natural, semejantes a un magnetofón de bolsillo, que permiten obtener varias respuestas psicofisiológicas.

Diversos autores han analizado varios problemas en la medición de las respuestas psicofisiológicas tales como: distorsiones provenientes del ambiente, o del organismo, reactividad y habituación, problemas en torno a la generalización de resultados, etc. Es importante tener en cuenta que ni la fiabilidad, ni la validez de la evaluación psicofisiológica han sido establecidas firmemente por los criterios tradicionales (Matud, 1993d).

Respecto a la contribución de las variables psicofisiológicas a la evaluación clínica, Haynes (1991) destaca que aportan información sobre la operatividad de procesos psicológicos importantes, pueden funcionar como factores de riesgo para algunas alteraciones conductuales y pueden ser útiles en el diseño de programas de intervención clínica y como medidas del resultado del proceso terapéutico. En la misma dirección, Ozer y Reise (1994) consideran que la evaluación psicofisiológica puede complementar útilmente la investigación de la personalidad proveyendo un importante método alternativo de medida, especialmente valioso cuando los autoinformes puedan ser de cuestionable validez o los juicios de los observadores sean relativamente dudosos.

Así, la evaluación psicofisiológica requiere una instrumentación específica para la presentación de los estímulos: generadores y programadores de estímulos auditivos o visuales, y ordenadores, entre otros equipos. Por otro lado, las tareas que debe realizar el sujeto están en función del aspecto psicológico que se pretende estudiar. Varían si lo que se estudia es, por ejemplo, el estrés, las emociones, o las respuestas atencionales, aunque se registren las mismas variables fisiológicas. Algunos ejemplos de tareas son: hablar en público, resolución de tareas aritméticas, apretar un botón cuando aparezca un estímulo determinado, hinchar un globo hasta que explote, dibujar una figura en espejo, o simplemente permanecer en reposo ante una serie de estímulos auditivos o visuales (Fuentes y Martínez, 2003).

Respecto a las aplicaciones de las técnicas psicofisiológicas existen numerosos campos de aplicación de estas técnicas: psicología clínica, medicina conductual, neuropsicología,

psicología social, ergonomía, deporte, detección del engaño.... En general, se pueden destacar 4 áreas fundamentales de aplicación (Calero y Fernández-Ballesteros, 2004).

1. La evaluación psicofisiológica de los trastornos psicopatológicos ha sido una de las áreas que ha experimentado mayor desarrollo en los últimos años; gran parte de este desarrollo ha estado orientado a la búsqueda de los llamados marcadores psicofisiológicos. Los marcadores psicofisiológicos son desviaciones en una variable fisiológica que están asociadas de forma fiable con un trastorno en particular. Existen diferentes tipos de marcadores: los marcadores episódicos, los marcadores de vulnerabilidad y los marcadores genéticos. Los marcadores episódicos indican el inicio y el final de los síntomas del trastorno. Los marcadores de vulnerabilidad indican la existencia de una mayor probabilidad de desarrollar el trastorno. Finalmente, los marcadores genéticos indican la presencia de uno o varios genes que explican la mayor vulnerabilidad a desarrollar la patología. Aportaciones importantes en este sentido se están realizando en determinados trastornos psicopatológicos; en particular, en los trastornos de la ansiedad, la esquizofrenia, la depresión y la psicopatía (Vila, 2004).
2. En el ámbito de los trastornos psicofisiológicos, la evaluación psicofisiológica se utiliza a dos niveles: para establecer un diagnóstico diferencial y para valorar la intervención psicoterapéutica. Entre los principales trastornos en los que la respuesta psicofisiológica es importante para ser considerada se encuentran: trastornos cardiovasculares, como la hipertensión esencial, las arritmias cardíacas, la cardiopatía isquémica, la enfermedad de Raynaud y las cefaleas migrañosas; trastornos respiratorios como el asma bronquial y el síndrome de hiperventilación; trastornos gastrointestinales como la úlcera gástrica o duodenal, el colon irritable, la colitis ulcerosa y el estreñimiento crónico; trastornos dermatológicos como el eccema, el acné, la soriasis y la alopecia; trastornos sexuales como las disfunciones del deseo, la excitación y el orgasmo; trastornos neuromusculares como los tics, los temblores, las lumbalgias y las cefaleas tensionales; y trastornos neurológicos centrales como determinados trastornos del sueño y la epilepsia. Por otra parte, dada la importancia que las técnicas de relajación y las técnicas de biofeedback tienen en el tratamiento psicológico de estos trastornos, la evaluación pre y postratamiento de las variables fisiológicas objeto del entrenamiento en relajación o biofeedback debería convertirse igualmente en parte integrante del protocolo de intervención (Vila, 2004).
3. La evaluación mediante técnicas objetivas en general y mediante registros en particular también desempeña un papel importante en el área de los trastornos neuropsicológicos. Los potenciales evocados fueron los primeros tipos de registro utilizados en trastornos como la enfermedad de Alzheimer, la epilepsia, los trastornos motores, el alcoholismo y el retraso mental (Simón y Amenedo, 2001). Aunque en menor medida, la cartografía de la actividad eléctrica cerebral y las técnicas metabólicas de neuroimágenes también han empezado a ser utilizadas en este ámbito.
4. La evaluación en psicología jurídica en relación con la detección del engaño.

7.7.3. VALIDEZ Y FIABILIDAD DE LAS TÉCNICAS OBJETIVAS

Las técnicas objetivas, por la estructuración de su material y las condiciones de control de laboratorio con que son aplicadas, suelen presentar una apreciable validez interna de sus datos, sin embargo, suelen tener dificultades en cuanto a su validez externa y ecológica. Los sujetos no pueden manipular sus respuestas a voluntad siendo esto una ventaja, sin embargo, la situación experimental suele provocar reactividad en el comportamiento de los sujetos. Con esta preocupación en su base algunos estudios se están focalizando en el análisis de las condiciones ambientales más apropiadas para la evaluación psicofisiológica.

El hecho de que el material de prueba sea tan sensible y, a veces de difícil manejo, puede promover múltiples sesgos tanto provenientes de factores externos como del evaluador que pueden influir negativamente en una toma de datos fiable. Si bien una ventaja de estos aparatos es el hecho de que producen puntuaciones objetivas, hay autores que llaman la atención sobre la diferencial forma de interpretar estas puntuaciones. De nuevo emerge la polémica entre utilizar las puntuaciones como indicadores o índices de construcciones internas, es decir, como signo de entidades intrapsíquicas, o bien hacerlo como muestra del comportamiento del sujeto.

En torno a la medición de las respuestas psicofisiológicas se han destacado, principalmente, los siguientes problemas (Fernández-Ballesteros y Calero, 1992; Matud, 1993d):

- *Artefactos*: El principal problema que tienen los registros psicofisiológicos estriba en la posibilidad de atribuir los cambios observados en la variable fisiológica (dependiente) objeto de estudio como una respuesta al estímulo al que el sujeto ha sido expuesto. Carrobbles (1981) clasifica los artefactos en función de su fuente o procedencia: artefactos ambientales y de registro (campo eléctrico, luz, temperatura, ruido...) y artefactos procedentes del organismo (movimientos del sujeto).
- *Ley de los valores iniciales*: Las relaciones entre los estímulos ambientales y una determinada respuesta psicofisiológica depende del nivel previo de activación de esa respuesta. Así, cualquier respuesta fisiológica frente a un estímulo estará en dependencia no sólo de éste, sino también de los niveles previos de esa respuesta.
- *Reactividad y Habitación*: La situación experimental en que se recogen los registros psicofisiológicos provoca comportamientos específicos no relacionados con la estimulación que le es presentada sino con todo el conjunto de estímulos de la situación a examen. La única manera de controlar los efectos de la reactividad del sujeto es a través de un período de habituación por medio de sesiones pre-experimentales que tengan por finalidad conseguir la adaptación del sujeto a la situación experimental.
- *Balance Autónomo y Homeostasis*: Otra variable que afecta a los registros psicofisiológicos es el nivel homeostático y proceso de autorregulación que el organismo ejerce sobre sus respuestas.
- *Problemas en torno a la generalización de los resultados*: La situación artificial, de laboratorio difiere de la situación en la vida real, por lo que habrá que ser cautelosos a la hora de generalizar los resultados de un ámbito a otro.

En la evaluación psicofisiológica, el contexto psicológico del registro es el que confiere a la variable fisiológica registrada y analizada el valor de índice psicofisiológico, esto es, la capacidad de convertirse en indicador de procesos psicológicos o conductuales (Vila, 1994). De ahí que el contexto del registro sea tan importante o más que los aspectos técnicos y metodológicos relativos al polígrafo y al sistema de computerización. En opinión de Vila (1994) el contexto del registro incluye tres aspectos importantes a considerar: el laboratorio, el sujeto experimental y las tareas experimentales.

7.8. LAS TÉCNICAS PSICOMÉTRICAS

7.8.1. DEFINICIÓN Y CARACTERÍSTICAS DE LOS TESTS PSICOMÉTRICOS

Hablar de técnicas psicométricas es hablar de tests psicológicos (Calero y Padilla, 2004). Tal y como ha sido definido por numerosos autores, un test psicológico tiene como características fundamentales:

- *Estandarización*: El material, instrucciones, procedimiento de aplicación e interpretación están estandarizados, es decir, son idénticos para todos los sujetos evaluados.
- *Corrección objetiva*: El resultado obtenido por el sujeto es independiente del profesional que realiza la evaluación, ya que el procedimiento de corrección está también estandarizado y no está sesgado por el conocimiento del evaluador.
- *Resultado cuantitativo*: Los resultados ofrecen un resultado numérico.
- *Tipificación*: La interpretación sólo tiene sentido a partir de la comparación de las puntuaciones obtenidas por el sujeto con un grupo de referencia normativo, es decir, con un grupo que se establece como norma.

A pesar de las críticas recibidas y de que se han ido elaborando otros tipos de técnicas de evaluación, el uso de los tests psicológicos sigue estando muy extendido cuando lo que se pretende es la contrastación de hipótesis a nivel correlacional, o cuando el objetivo del proceso es la descripción, selección y/o la orientación del sujeto, cuando se trata de determinar la posición del sujeto frente a otros o evaluar el grado en que un sujeto presenta determinado repertorio conductual (Calero y Padilla, 2004).

Los orígenes y primeros desarrollos de la evaluación psicológica estuvieron íntimamente ligados al desarrollo de los tests; posteriormente se fueron desarrollando nuevos y diferentes procedimientos de evaluación para diversos objetivos, y esto, junto a una postura crítica sobre el tipo de evaluación que los tests representaban, provocó, en cierta medida, un declive en su uso y un freno a su desarrollo. No obstante, aunque ahora de modo menos exclusivo, los tests psicológicos siguen desempeñando un papel importante en la psicología aplicada.

7.8.2. CLASIFICACIÓN DE LOS TESTS PSICOMÉTRICOS

La gran cantidad de tests psicológicos publicados a lo largo de la historia de la evaluación psicológica ha llevado a muchos autores a intentar establecer algún tipo de clasificación. La mayoría de las clasificaciones que se hacen de los tests son similares a las que se aplican para las técnicas de evaluación en general y se estructuran con referencia a un criterio único. Por ello, como observan Calero y Padilla (2004), podemos encontrar clasificaciones en función de

criterios de contenido, que se refieren a las funciones psicológicas que el test pretende evaluar; se dividen entonces en: tests de inteligencia, tests de personalidad...; o en función de criterios materiales, estableciendo como grupos: tests manipulativos, de papel y lápiz...; también se utilizan criterios basados en la población a la que se dirigen, tales como: infantiles, de adultos, etc., y más frecuentemente, el criterio de clasificación utilizado se centra en los modos de aplicación: individuales, colectivos, etc.

Por otro lado, hay 2 clasificaciones adecuadas de las técnicas psicométricas. La primera hace referencia a las estrategias de construcción seguidas para la elaboración del test: a) teóricos, racionales, empíricos; y/o b) analítico/factoriales. La segunda clasificación realizada por Cronbach (1972), según la demanda de rendimiento que se le hace al sujeto: a) tests de rendimiento máximo; y b) tests de rendimiento típico.

Los tests de rendimiento máximo son aquellos que quieren recoger los niveles más altos de rendimiento de una persona, aquellos que exigen al sujeto que rinda al máximo de sus posibilidades (tests de inteligencia y/o aptitudes). Los tests de rendimiento máximo tienen una única respuesta verdadera, y ésta es la que se requiere del sujeto; lo importante en cada ítem es su nivel de dificultad o la rapidez de ejecución que exige. Según esto la clasificación dentro del grupo se establece entre tests de potencia y tests de velocidad. En los tests de potencia, los ítems van aumentando progresivamente de dificultad, y se trata de determinar el nivel máximo de ejecución que un sujeto alcanza sin límite de tiempo; esto es, el nivel de habilidad de que el sujeto es capaz. Los tests de velocidad se componen de ítems con índices de dificultad similares y en los que lo importante es determinar la cantidad de ítems bien resueltos en un tiempo limitado, porque lo que interesa establecer al evaluador es el nivel de eficacia y precisión, a ritmo impuesto, que una persona posee.

El grupo de tests de rendimiento típico estaría formado por todos aquellos instrumentos que intentan recoger el modo habitual de comportarse de un individuo, o sea, tests de personalidad y cuestionarios, escalas o inventarios de intereses y/o actitudes. Volviendo a Pervin, este grupo sí poseería la característica de ser técnicas voluntarias, pues, en estos tests, el sujeto percibe, generalmente, el constructo que se trata de evaluar y, por tanto, tiene el control sobre sus respuestas. Además, en muchas ocasiones estos tests se interesan por recoger la opinión o juicio de la persona evaluada, no existiendo por tanto una respuesta verdadera, sino diferentes opciones de respuesta con distinto significado psicológico.

En la línea de la clasificación realizada por Cronbach, en este proyecto se diferencian dos grandes grupos de tests psicológicos: los tests de inteligencia y los tests de personalidad.

7.8.2.1. TESTS DE INTELIGENCIA: TESTS INDIVIDUALES Y COLECTIVOS

Los tests de inteligencia pueden clasificarse en dos grandes grupos: los que necesariamente obligan a realizar una aplicación individual, y los que pueden ser aplicados de forma colectiva. Además, dentro de estos se pueden diferenciar los que evalúan un factor de inteligencia general y los que exploran varias aptitudes intelectuales. Un resumen con algunos de los tests de inteligencia que se utilizan con bastante frecuencia puede observarse en el Cuadro 29.

Cuadro 29. Tests de inteligencia: tests individuales y colectivos

Clasificación de los Tests	Denominación del Test y Autor	Aplicación	Significación de la prueba
Tests de Inteligencia Individuales	Wechsler, D. (1999). WAIS III: Escala de inteligencia de Wechsler para adultos. (original, 1997).	16-94 años	Inteligencia General: Inteligencia Verbal - Inteligencia Manipulativa
	Wechsler, D. (2005). WISC-IV: Escala de Inteligencia de Wechsler para niños-revisada. (original, 1949/2003).	6-16 años	
	Wechsler, D. (1976). WPPSI: Escala de Inteligencia de Wechsler para preescolar y primaria. (original, 1976).	4-6 años	
	Kaufman, A.S. y Kaufman, A.L. (1997). K-ABC: Bateria de Evaluación de Kaufman para niños. (original, 1983).	2,5-12,5 años	Inteligencia General: Procesamiento simultáneo - Procesamiento sucesivo - Conocimientos.
Tests de Inteligencia Colectivos	Kaufman, A.S. y Kaufman, N.L. (1997). K-bit: Test breve de inteligencia. (original, 1990).	4-90 años	Inteligencia General: Inteligencia Verbal - Inteligencia No Verbal
	Raven, J.C., Court, J.H., y Raven, J. (1994). Matrices Progressivas de Raven. (original 1938).	5 años en adelante	Inteligencia General No Verbal - Razonamiento Analógico.
	• Escala Color. CPM. Preferentemente niños de 5,6 a 11 años y personas mayores con deficiencias mentales o deterioro. • Escala General. SPM. Preferentemente para adolescentes y adultos, aunque el manual presenta baremos para niños desde 9 años. • Escala Superior o Avanzada. APM. Para sujetos con un nivel de inteligencia superior, sujetos universitarios.		
	Tests de Dominós: D-48 (Anstey, 1944/1979), D-70 (Kowrousky y Remes, 1970/1980) TIG-1 y TIG-2 (TEA, 1971-1973/1995).	A partir de 12 años 10 años	Inteligencia General No Verbal (Abstracción y comprensión de relaciones)
Tests de Aptitudes	TONI-2. Test de Inteligencia No Verbal (Brown, Sherbenou, y Johnsen, 1990/1995).	De 5 a 85 años	Inteligencia General No Verbal (Razonamiento abstracto)
	TRASI. Test informatizado para la evaluación del Razonamiento secuencial y la inducción. (Rubio y Santacreu, 2003)	Adultos	Inteligencia General No verbal utilizando el razonamiento y la aptitud numérica. Aplicación computerizada.
	PMA. Aptitudes Mentales Primarias (Thurstone, 1939/2007).	A partir de 10 años	Comprensión Verbal. Aptitud Espacial. Razonamiento. Calculo Numérico. Fluidez Verbal.
	DAT-5. Test de Aptitudes Diferenciales (Bennet, Seashore y Wesman, 2000).	A partir de 12 años	Razonamiento: Verbal. Numérico. Abstracto. Rapidez y exactitud perceptiva. Razonamiento Mecánico. Relaciones Espaciales. Ortografía.

Tests de inteligencia individuales

A finales del siglo pasado, las primeras concepciones de la inteligencia hacían referencia a la capacidad motora y sensorial del sujeto, y con esta perspectiva se elaboraron pruebas sensoriales y motoras para medirla. Sin embargo, a principios de siglo XX, Binet (Binet y Simon, 1905) supera esta concepción de la inteligencia y la concibe como una capacidad mental global, como un conjunto de funciones cognitivas (comprensión, razonamiento, imaginación...), empezando a considerar la inteligencia como una entidad compuesta.

La evaluación de la inteligencia comienza con la publicación de la “escala de inteligencia” de Binet y Simon en 1905 y en cierto modo la mayoría de los instrumentos de evaluación de la inteligencia se han elaborado como desarrollo o como alternativa a esta escala. La estrategia de construcción seguida por Binet y Simon fue una estrategia racional, pues se trataba de evaluar mediante observación un conjunto de funciones cognitivas que sus autores consideraban como demostrativas o relacionadas con la inteligencia. El concepto implícito de inteligencia era el de capacidad global, y el test desarrollado estaba formado por diferentes tareas (que se supone requieren una habilidad global) seleccionadas por orden de dificultad creciente y agrupadas según criterios de edad. En sus primeras versiones la escala de Binet medía la inteligencia en términos de edad mental (asignación de una puntuación, expresada en unidades de edad, por comparación con las tareas que un niño normal de esa edad determinada ejecuta por término medio); posteriormente éste pasó a CI (Cociente Intelectual) al dividir la EM por la edad cronológica. Mucho más tarde, Weschler introduce el CI de desviación, o sea, lo transforma en una puntuación típica normalizada con una media de 100 y desviación típica de 15, y ésta es la unidad de medida de esa capacidad general que actualmente se sigue utilizando y que adoptan la mayoría de los tests individuales de inteligencia general.

El objetivo de la evaluación desarrollada a partir del test de Binet era la de alcanzar una puntuación indicativa del nivel global de inteligencia del sujeto (CI) mediante la ejecución individual de diferentes tareas. Con el tiempo este modo de evaluación fue desarrollado en diversos instrumentos, mejorados por las contribuciones teóricas pero que en esencia mantienen los supuestos y modos de evaluación primitivos. Los test de inteligencia individual tienen varias diferencias con los tests colectivos: 1) son de aplicación más larga, requieren un mayor adiestramiento del psicólogo que los administra pues se trata de aplicar tareas muy diferentes, algunas de ellas con control de tiempo, otras con manipulación de materiales; 2) proporcionan mayor información ya que suelen ofrecer varios tipos de puntuaciones, perfiles de habilidades...; y 3) permiten la observación del sujeto durante su ejecución, por lo que se consideran más útiles desde el punto de vista clínico.

Entre los principales tests individuales de inteligencia adaptados y baremados con población española y que actualmente se utilizan en España caben destacar: las Escalas de Wechsler “la WAIS III para adultos (Wechsler, 1999), la WISC-IV (Wechsler, 1949/2005) para niños, la WIPPSI (Wechsler, 1963/1976) para preescolares” y las Escalas Kaufman “K-ABC, para niños (Kaufman y Kaufman, 1993/1997) y la escala breve K-BIT para niños, adolescentes y adultos (Kaufman y Kaufman, 1990/1997). Tanto las escalas de Wechsler como las de Kaufman comparten un planteamiento teórico similar, que siguiendo a Wechsler (1975) se puede conceptualizar del siguiente modo: “lo que miden los tests (individuales) de inteligencia es la capacidad del sujeto de comprender el mundo que le rodea y los recursos que posee para enfrentarse con sus exigencias y desafíos”. Por lo tanto la inteligencia se entiende como un

conjunto de habilidades o capacidades que permiten la adaptación del sujeto a las exigencias del medio, y esto hace que estén compuestos por tareas muy diversas. Las pruebas recogen datos sobre la ejecución actual de la persona, considerando que ésta es estable, que puede predecir futuras ejecuciones y por ello se utilizan para clasificación y/o diagnóstico.

Las escalas de inteligencia de Wechsler contienen un conjunto de tareas verbales y no verbales y permiten obtener un CI verbal, un CI no verbal y del sumatorio de ambos un CI global. Estas escalas incorporaron el CI de desviación, que es una medida en puntuaciones típicas normalizadas de la inteligencia. Las 3 escalas de inteligencia de Wechsler permiten apreciar “la Inteligencia Verbal” (CIV), es decir, la adquisición y utilización del lenguaje, la habilidad cognitiva verbal, “la Inteligencia Manipulativa o de Ejecución” (CIM) mediante tareas de tipo perceptivo y manipulativo, habilidad cognitiva no verbal, así como “la Inteligencia Total” (CIT) o capacidad cognitiva general. Además, la WAIS III permite apreciar otras agrupaciones basadas en unos aspectos más precisos del funcionamiento cognitivo, que dan lugar a los índices “Comprensión verbal” (CV), “Organización perceptiva” (OP), “Memoria de trabajo” (MT), y “Velocidad de Proceso” (VP).

Las escalas de Kaufman, K-BIT (Kaufman y Kaufman, 1990/1997) y K-ABC (Kaufman y Kaufman, 1983/1997), son instrumentos clínicos de aplicación individual, diseñadas en la década de los 80 y al inicio de los 90, las cuales incorporan los avances recientes en la elaboración de pruebas, y representan un intento de alejarse de la posición ateorica de las anteriores escalas de inteligencia, de tal modo que pudieran afianzarse en las teorías en evolución sobre la inteligencia. De acuerdo con las aportaciones de la neurología y la psicología cognitiva, los autores conciben la inteligencia en términos del estilo individual de procesar la información, como “la habilidad para resolver problemas mediante procesos mentales de carácter simultáneo y secuencial”. Como señalan Forns y Amador (1995b) estas escalas se basan en la convergencia de diferentes modelos cognitivos y se fundamentan en una concepción dicotómica del procesamiento de la información que supone, por un lado, habilidades analíticas, secuenciales, sucesivas o seriales y, por otro lado, habilidades holísticas, gestálticas, simultáneas o paralelas.

Las Escalas de Kaufman proporcionan una medida de la inteligencia entendida por sus autores como estilo individual de resolver problemas y procesar informaciones. En su construcción incorporan las teorías más recientes (procesamiento de la información) sobre la inteligencia como base de construcción de sus pruebas. Permiten evaluar el tipo de procesamiento que un individuo utiliza en mayor grado para resolver un problema. Por ello, diferencian entre tareas de procesamiento simultáneo, tareas de procesamiento sucesivo y conocimientos, en el K-ABC-test que está adaptado al castellano-, y añaden, además, otros subtests dirigidos a evaluar las operaciones formales según el esquema teórico de Piaget y habilidades de planificación, memoria a largo plazo y estado mental, según el esquema teórico de Luria en el KAIT (escala que aún no está adaptada a nuestro país). En esta última, la parte fluida presenta mayores diferencias con las Escalas Wechsler que el K-ABC. El área de procesamiento secuencial se basa en la seriación o el orden temporal de presentación de los estímulos. El área de procesamiento simultáneo requiere una representación grafica, frecuentemente espacial, así como la integración eficaz de estímulos para resolver un problema. Estas dos áreas y su combinación (procesamiento mental compuesto) se consideran próximas a la inteligencia fluida. La escala de conocimientos incluye información novedosa seleccionada en función de criterios lógicos y racionales para medir aquellas habilidades

tradicionalmente recogidas en tests de inteligencia general de tipo verbal. Esta escala se considera próxima a la inteligencia cristalizada. La importancia de estas escalas está en que la clasificación de tareas se basa en tipos de procesamiento en vez de en contenidos (verbales/no verbales), como ocurría en las escalas de Wechsler, y, por tanto, no evalúan exactamente los mismos aspectos que las escalas de Wechsler aunque compartan algunos factores con éstas, sobre todo con respecto a los subtests cristalizados.

Tests de inteligencia colectivos: inteligencia general frente a batería de aptitudes

Los tests de inteligencia colectivos se inician con los tests Alfa y Beta Army (Anastasi y Urbina, 1998). Como han señalado Calero y Padilla (2004), a nivel práctico, los tests colectivos: 1) implican un mayor grado de estructuración que los individuales, pues requieren respuestas cerradas de opción múltiple; 2) en la aplicación se deja al sujeto sólo frente a la tarea hasta que termina la tarea o el tiempo concedido para realizarla, y 3) en la aplicación se tiene un escaso control sobre el grado de comprensión que tienen los sujetos de las instrucciones. Sin embargo, su uso ha sido muy frecuente y se ha impuesto en determinados contextos (selección profesional, educativos...) debido que son más económicos, porque la aplicación y corrección suele ser más rápida, permiten aplicaciones masivas y exigen menor formación del evaluador.

A nivel teórico los estudios de análisis factorial realizados en los años 20 y 30 del pasado siglo sobre estos tests de inteligencia, influyeron en el desarrollo de dos posturas teóricas sobre este constructo: las teorías monofactoriales, que mantenían la existencia una única capacidad global, frente a las posturas multifactoriales, que, frente a esa única habilidad general, defendían la existencia de múltiples aptitudes independientes o factores específicos relacionados con diferentes tipos de tareas.

Algunos investigadores identificaron a esa inteligencia global con el factor general de inteligencia: factor g, que se relacionaba con tareas no verbales de razonamiento inductivo (por ejemplo, las matrices progresivas de Raven, 1938). El concepto factor cognitivo general, factor "g" fue propuesto por Spearman en 1927 (Spearman, 1927, 1936). Este psicólogo matemático llegó a la conclusión de que el proceso psicológico que subyace a los resultados que reflejan la existencia del factor "g" es un proceso noegenético y de abstracción. Entendiendo por noegenético tres aspectos de la vida mental de un individuo: (1) la capacidad de introspección, de observar lo que ocurre dentro de su mente; (2) la capacidad de educir las relaciones esenciales existentes entre dos o más ideas que se tienen en mente; y (3) la capacidad que una persona tiene, supuesto que tenga en su mente una idea y una relación, de concebir la idea inicial implícita en esa relación. El elemento crucial responsable del grado en que un test satura en "g" es la cantidad de manipulación mental consciente desencadenada por la información presentada al sujeto. Se trata de una habilidad fundamental que interviene en todas las operaciones mentales, representa la energía mental y se moviliza en toda tarea no automatizada. Los tests de matrices fueron diseñados para evaluar el factor "g" tal como fue definido por Spearman.

Sin embargo, otros investigadores llegan a aislar diversos factores que denominan aptitudes mentales (comprensión verbal, fluidez verbal, razonamiento numérico, memoria...) lo que da lugar a los tests de aptitudes mentales (por ejemplo, PMA de Thurstone 1938/2007).

Algunas reflexiones sobre el cociente intelectual y la evaluación de la inteligencia

El cociente intelectual o CI se utiliza como designación abreviada de la inteligencia y representa la puntuación obtenida en un test. En la misma dirección que Anastasi y Urbina (1998) se pueden realizar algunas precisiones en torno a la inteligencia y al significado del CI:

1) *El CI es un concepto descriptivo y no explicativo.* El CI describe el desempeño de un sujeto pero no explica las causas del mismo. Las pruebas psicológicas en general y los tests de inteligencia en particular no deben utilizarse para clasificar a los sujetos sin más, sino para comprender a las personas y ayudarlas a desarrollar su nivel óptimo de funcionamiento mediante la planificación de programas de entrenamiento cuando sea necesario.

2) *La inteligencia no es una capacidad unitaria sino un compuesto de varias funciones.* El término suele emplearse para cubrir la combinación de habilidades que una cultura requiere para sobrevivir y progresar, por lo que es un constructo que se modifica en el tiempo. Los tests tradicionales de inteligencia miden principalmente las habilidades que predominan en el aprendizaje escolar, por lo que en gran medida pueden considerarse instrumentos de evaluación de la inteligencia académica. Por ello, el CI puede ser una medida del aprovechamiento educativo previo y un predictor del rendimiento académico futuro. Puesto que, como se ha podido observar previamente, cada test de inteligencia explora distintas funciones y procesos intelectuales es necesario especificar el test con el que se ha obtenido ese CI para comprender realmente sus implicaciones. Cabe en última instancia resaltar que hay muchas funciones psicológicas importantes que los tests de inteligencia no miden, en este sentido llama la atención la significativa ausencia de tareas que evalúen la creatividad y sus indicadores (flexibilidad de pensamiento, originalidad, conectividad...), siendo un factor cognitivo de gran relevancia en el proceso de adaptación del hombre a la realidad.

3) *Estabilidad – Inestabilidad del CI o de la ejecución en los tests de inteligencia.* Existen estudios que evidencian conclusiones en ambas direcciones: a) Un grupo de estudios ha mostrado la estabilidad creciente del CI, es decir, que las ejecuciones en los tests de inteligencia permanecen relativamente constantes a lo largo de la escolarización. No obstante, esta es una estabilidad creciente debido a la naturaleza acumulativa de la inteligencia ya que con la edad el individuo incorpora nuevos conocimientos y habilidades...; y b) Diversas observaciones también muestran que el CI no es algo fijo e inmutable sino que puede ser modificado por el ambiente. Estudios evidencian aumentos o disminuciones en las puntuaciones que un niño tiene en las pruebas de inteligencia que pueden estar relacionadas con cambios ambientales significativos en la vida de ese niño, enfermedades u otros eventos que concurran en su vida. Esta segunda línea de estudios ha dado pie al desarrollo de programas de entrenamiento para la mejora de la inteligencia. Desde finales de la década de los 70 se han desarrollado programas de entrenamiento de habilidades cognitivas (resolución de problemas, lenguaje, memoria, pensamiento analógico...) que han puesto de relieve el importante papel que puede desempeñar el ambiente en cambios que se operan en la misma. Estos resultados han cuestionado anteriores concepciones de la inteligencia que desde una perspectiva de herencia genética la habían considerado como inmodificable.

Aunque en los últimos años se han realizado importantes avances en la investigación de la inteligencia y su evaluación, los tests de inteligencia tradicionales no están exentos de problemas entre los que se pueden destacar:

1) Un problema conceptual ya que el constructo "inteligencia" no está claramente definido, a este término se le atribuyen diferentes significados y se requiere una definición precisa y aceptada unánimemente. Aunque en el ámbito de la investigación se ha evolucionado en una definición más compleja de la inteligencia que incorpora las aportaciones de la psicología cognitiva, y desde la que se define la inteligencia en términos de estrategias, conocimientos, resolución de problemas, componentes, experiencia, esquemas..., sin embargo, esta modificación en la conceptualización no ha tenido su total correlato en la práctica aplicada. Cabe resaltar que un intento por integrar las teorías del procesamiento de la información con la psicometría tradicional lo representan las escalas de Kaufman, y que en el futuro queda por integrar las contribuciones de la investigación neurológica, bioquímica y fisiológica relacionada con la actividad inteligente.

2) La influencia cultural de los tests de inteligencia ya que los sujetos con baja escolarización obtienen puntuaciones inferiores y en muchas ocasiones se utilizan los tests sin ser adaptados a las características de esa población específica. Como observan Almeida y Buela-Casal (1997) los tests de inteligencia tuvieron un gran auge en los años 20 y 30 especialmente en Estados Unidos, pero progresivamente, en la segunda parte del S. XX surgió una importante controversia social debida a las puntuaciones inferiores que presentaban las minorías étnicas con el impacto subsiguiente en la posibilidad de ocupar buenas escuelas o empleos, lo que evidenciaba su ausencia de neutralidad social y validez científica. Los tests de inteligencia libres de influencia cultural surgen como un intento para evitar la influencia de los conocimientos, la escolarización y la cultura en la evaluación de la inteligencia. Otra dirección que se plantea para resolver este problema es la de utilizar tests criterioles, tests que se centran más en el análisis intraindividual que en el análisis comparativo desde las normas de grupo de los tests normativos actuales. Por otro lado, asumiendo que ningún test está libre de influencia cultural Sternberg (1985ab, 1991) ha desarrollado el STAT (Sternberg Triarchic Abilities Test) diseñado para evaluar procesos y funciones de las tres subteorías que componen la teoría triárquica (componencial, experiencial y contextual), lo que permite tener en cuenta el contexto y la experiencia del individuo.

3) La existencia de diversos factores o efectos que ejercen una influencia significativa en los resultados de los tests de inteligencia como son: a) el efecto del entrenamiento en la mejora de los resultados; b) el efecto de las expectativas de los resultados por parte del sujeto o de otros que pueden afectar significativamente (efecto Rosenthal); c) el efecto del evaluador y de las condiciones de la aplicación del test (espacio físico donde se realiza la evaluación, actitud del evaluador en el proceso, interacción entre ambos...); d) el efecto de características de personalidad tales como motivación de logro, persistencia, sistema de valores... la relación entre personalidad e inteligencia es recíproca ya que las características de la personalidad influyen en el desarrollo intelectual y el nivel de inteligencia puede afectar al desarrollo de la personalidad; y e) el efecto de estados afectivos, la existencia o ausencia de problemas emocionales...

7.8.2.2. TESTS DE PERSONALIDAD: PRUEBAS MULTIRASGO Y MONORASGO

Existen numerosas definiciones diferentes de personalidad en la literatura psicológica. Una definición reciente que aborda esta cuestión en una perspectiva integradora es la de Cohen y Swerdlik (2001) que después de realizar una amplia revisión de diferentes definiciones del constructo definen la personalidad como una constelación única de rasgos y estados psicológicos del individuo. En esta definición se incluyen valores, intereses, actitudes, perspectiva del mundo, aculturación, identidad personal, humor, estilos cognoscitivos y conductuales; y desde esta conceptualización, se entiende la evaluación de la personalidad como la medida de rasgos psicológicos, estados, valores, intereses, actitudes, perspectiva del mundo, aculturación, identidad personal, sentido del humor, estilos cognoscitivos y conductuales, y/o características individuales relacionadas.

La evaluación de la personalidad mediante tests aparece un poco más tarde que la evaluación de la inteligencia, lo que implica que la teoría psicométrica clásica estaba más desarrollada y supone que los diferentes instrumentos de evaluación de la personalidad se elaboren siguiendo una de las diferentes estrategias de construcción comentadas en función de su base teórica y de sus objetivos (Calero y Padilla, 2004). Las técnicas psicométricas de personalidad, esto es los tests de personalidad surgen dentro de un modelo de rasgos. Desde esta perspectiva se asume que personalidad es lo que permite predecir lo que una persona hará en determinada situación (Cattell, 1950). En definitiva, se supone que el comportamiento humano puede ser ordenado y medido según unas dimensiones relativamente estables denominadas rasgos que lo caracterizan razonablemente bien.

Respondiendo a esa conceptualización se han desarrollado una serie de tests o cuestionarios multidimensionales, es decir, que evalúan un conjunto de rasgos de personalidad, mientras que otros se han centrado en un único rasgo, en algún rasgo específico. Algunos ejemplos de los test psicométricos se utilizan con mayor frecuencia para evaluar rasgos de personalidad pueden observarse en el Cuadro 30.

Cuadro 30. Tests de personalidad: pruebas multirasgo y monorasgo

Denominación del test	Significación	Edades
MMPI-2 Inventario multifásico de personalidad de Minnesota. (Hathaway y Mckinley, 1999)	10 Escalas clínicas: hipocondría, depresión, histeria, desviación psicopática, masculinidad-feminidad, paranoia, psicastenia, esquizofrenia, hipomanía, introversión social. 15 Escalas de contenido: ansiedad, miedos, obsesividad, depresión, preocupación por la salud, pensamiento extravagante, hostilidad, cinismo, conductas antisociales, comportamiento tipo A, baja autoestima, malestar social, problemas familiares, interferencia laboral, indicadores negativos de tratamiento. 13 Escalas suplementarias: ansiedad, represión, fuerza del yo, alcoholismo, hostilidad controlada, dominancia, responsabilidad social, inadaptación universitaria, rol sobre el género, estrés postraumático, malestar matrimonial, potencial de adicción, reconocimiento de adicción.	19-65 años

16 PF Cuestionario factorial de personalidad (Cattell, 1972)	16 rasgos o factores de personalidad primarios: sizotimia-afectotimia; inteligencia baja-alta; poca-mucha fuerza del ego; sumisión-dominancia; desurgencia-surgencia; poca-mucha fuerza del super-yo; trectia-parmia; harria-premsia; alaxia-protensión; praxernia-autia; sencillez-astucia; adecuación imperturbable-tendencia a la culpa; conservadurismo-radicalismo; adhesión al grupo-autosuficiencia; baja integración-mucho control de la autoimagen; poca tensión energética-mucha tensión energética. 4 factores de personalidad secundarios: ansiedad baja-alta; introversión-extraversión; poca-mucha socialización controlada; dependencia-independencia.	De 16 años en adelante
16 PF-5 Cuestionario factorial de personalidad 5ª edición (Cattell, Cattell y Cattell 1993)	16 factores o rasgos primarios de personalidad: afabilidad, razonamiento, estabilidad, dominancia, animación, atención-normas, atrevimiento, sensibilidad, vigilancia, abstracción, privacidad, aprensión, apertura-cambio, autosuficiencia, perfeccionismo, tensión. 5 dimensiones globales: extraversión, ansiedad, dureza, independencia y autocontrol.	Adolescentes y adultos
HSPQ Cuestionario de personalidad para adolescentes (Cattell y Cattell, 1968)	14 rasgos de personalidad primarios: reservado-abierto, inteligencia baja-alta, emocionalmente afectado-estable, calmoso-excitante, sumiso-dominante, sobrio-entusiasta, despreocupado-consciente, cohibido-emprendedor, sensibilidad dura-blanda, seguro-dubitativo, sereno-aprensivo, sociable-autosuficiente, menos-más integrado, relajado-tenso. 4 rasgos secundarios: ajuste-ansiedad, introversión-extraversión, calma-excitabilidad, dependencia-independencia.	12-18 años
CPQ Cuestionario de Personalidad para niños (Porter y Cattell, 1959)	14 rasgos de personalidad primarios: los mismos que el HSPQ, con la excepción de sociable-autosuficiente que en el CPQ se sustituye por sencillo-astuto. 3 rasgos secundarios: ajuste-ansiedad, introversión-extraversión, patemia-excitabilidad/dureza.	8-12 años
ESPQ Cuestionario de Personalidad para niños (Coan y Cattell, 1966).	13 rasgos de personalidad primarios: los mismos que el CPQ, excepto menos-más integrado que no se incluye en el ESPQ. 2 rasgos secundarios: ajuste-ansiedad, introversión-extraversión.	6-8 años
BFQ Cuestionario Big Five. (Caprara, Barbaranelli y Borgogni, 1993/1995)	5 grandes dimensiones o factores de la personalidad y 10 subdimensiones: • Energía: dinamismo, dominancia. • Afabilidad: cooperación/empatía, cordialidad • Tesón: escurpulosidad, perseverancia • Estabilidad emocional: control de emociones, control de impulsos. • Apertura mental: apertura a la cultura, apertura a la experiencia	Adolescentes y Adultos

STAI. Cuestionario de Ansiedad Estado- Rasgo. (Spielberger, Gorsuch, y Lushene, 1970/1988)	Ansiedad-Estado Ansiedad-Rasgo	Adolescentes y Adultos
STAXI-2. Inventario de Expresión de Ira Estado/Rasgo (Spielberger, 2000)	Ira-Estado Ira-Rasgo Índice de Expresión de la Ira	Adolescentes y Adultos

7.8.3. PRESENTE Y FUTURO DE LOS TESTS PSICOMÉTRICOS

La encuesta aplicada por Muniz et al. (1999) a los psicólogos colegiados ha confirmado el amplio uso y el aprecio que tienen los profesionales por los tests. La utilización conjunta de tests y evaluación del rendimiento y su adecuado tratamiento psicométrico garantizan medidas de más calidad en las que basar las complicadas y en ocasiones trascendentes decisiones que tomamos con nuestras medidas.

En distintos momentos de han planteado diversas reticencias y críticas ante los tests y la medición. En relación a algunas críticas relativas al uso de los tests, Bravo y Fernández (2000) citan ocho: 1) Miden solo conocimiento declarativo y no procedimental; 2) Se centran en el resultado y no en el proceso; 3) No cubren adecuadamente el dominio evaluado; 4) Existen habilidades e inteligencias que no son evaluadas por los tests; 5) Son medidas normativas; 6) La estandarización supone una muestra homogénea, cuando la muestra evaluada es heterogénea, con diferente dominio del lenguaje, actitudes y valores; 7) Las calificaciones en los tests de opción múltiple dependen de atributos distintos de los que se quiere evaluar; 8) Al ser pruebas estandarizadas no miden las habilidades específicas de los contextos reales.

A partir de esta revisión crítica Bravo y Fernández (2000) concluyen que cada vez son más las razones que impulsan a los expertos en el campo de la medición a buscar alternativas. Bajo la precisión, objetividad y ahorro en tiempo y dinero que ofrece el uso de tests estandarizados, se realizan selecciones que reflejan sólo una o pocas dimensiones de las diferencias individuales, no se mide a toda la persona, cuando es precisamente toda la persona la que es contratada para un trabajo o seleccionada para realizar determinada tarea. Pues bien, en la misma dirección que otros autores (Garrido, Ponsoda, Abad y Olea, 2008) yo estaría de acuerdo con algunas de esas críticas. Con otras, estaría de acuerdo en que son correctas sólo si estamos ante tests mal contruidos. Por ejemplo, se pueden hacer ítems de opción múltiple que evalúen procesos mucho más complejos que el mero recuerdo. Con otras críticas, como la quinta (son medidas normativas), estoy en desacuerdo. Hemos visto que los tests pueden dar medidas normativas y referidas al criterio. Ahora bien, en lo que estoy totalmente de acuerdo es que lo deseable sería evaluar a “toda la persona”. La pregunta es cómo. Si es por la vía de procedimientos que no puedan probar su calidad métrica, entonces no; a pesar de que reciban el nombre de medición auténtica. Mi postura es si a la evaluación del rendimiento, si a los tests,

si a la entrevista estructurada... si a los distintos procedimientos que satisfagan los requisitos métricos, de forma que los valores registrados puedan considerarse medidas del atributo que queremos medir.

Como ya se ha comentado previamente en el capítulo sobre la evolución histórica de la evaluación psicológica, en diversos momentos se ha producido un cuestionamiento y crítica de los tests psicológicos, que parecía iba a derivar en un desuso de los mismos, sin embargo, los tests psicológicos siguen siendo hoy uno de los instrumentos más utilizados. Como han observado Calero y Padilla (2004), tres razones pueden explicar este hecho:

1. Uno de los objetivos de la evaluación psicológica es la selección profesional y/o educativa. Para este fin, es necesario utilizar instrumentos que posibiliten: a) la evaluación de gran número de sujetos, al mismo tiempo, en las mismas condiciones para todos ellos; y b) la comparación. Estos requisitos son los que guían la construcción de tests psicométricos colectivos, y por tanto, en este campo siguen siendo instrumentos de necesaria aplicación.
2. Cuando el objetivo de la evaluación psicológica es el diagnóstico, en muchas ocasiones se requiere contrastar correlacionalmente determinadas hipótesis antes de pasar a una evaluación más precisa del caso. Los tests psicológicos basan su interpretación en criterios normativos y por ello pueden ser útiles para este tipo de comprobación.
3. En un proceso de orientación escolar y/o laboral se puede solicitar al evaluador una descripción del individuo, y ello implica establecer perfiles de aptitudes y personalidad que ayuden a tomar decisiones. Muchos tests psicológicos en los últimos años se han orientado a la realización de perfiles comparativos.

Por consiguiente, el uso de tests psicológicos es adecuado para diversos objetivos de la evaluación, especialmente cuando se requiere la comparación del individuo con un grupo de referencia normativo. Por todo ello, se puede predecir que el uso de los tests psicológicos se perpetuará en el futuro, siempre que cumplan con las adecuadas garantías de fiabilidad y validez, y se construyan utilizando los nuevos desarrollos científicos. Como han señalado otros autores (Calero y Padilla, 2004) se pueden subrayar: 1) los beneficios considerables que están derivándose de la aplicación de modelos de procesamiento a los tests de inteligencia y aptitudes, suministrando, con ello una base teórica explicativa de los resultados que un individuo obtiene en la evaluación; 2) las ventajas, en cuanto a la precisión de las estimaciones, que se obtienen con la aplicación de la TRI en el diseño; 3) la economía que suponen los test adaptativos informatizados en cuanto a flexibilidad y posibilidad de evaluaciones masivas; y 4) el efecto beneficioso que el uso del ordenador está produciendo en esta metodología. Todo ello puede producir un mantenimiento e incluso un auge del uso de los tests psicológicos, siempre y cuando cumplan con las debidas garantías científicas.

7.9. TÉCNICAS DE RECOGIDA DE INFORMACIÓN EN LOS PROGRAMAS DE LA ASIGNATURA EVALUACIÓN PSICOLÓGICA EN LAS FACULTADES DE PSICOLOGÍA ESPAÑOLAS

Durante la primera quincena de octubre de 2008 se realizó una encuesta sobre contenidos básicos en torno a las técnicas que deben enseñarse en los programas de grado de "Evaluación Psicológica" en las Facultades de Psicología llevada a cabo por la profesora Carmen Almendros (UAM). Esta encuesta fue respondida por 18 profesores de la asignatura Evaluación Psicológica procedentes de diversas universidades españolas, con una media de años de experiencia docente en Evaluación Psicológica de 18,5 años (DT: 10,37) en un rango entre 1 y 40 años. Los resultados sobre la docencia en técnicas de evaluación psicológica han puesto de relieve los datos que se presentan en los Cuadros 31 y 32.

Cuadro 31. Técnicas de recogida de información a incluir en el programa docente de la asignatura Evaluación Psicológica desde la opinión de los docentes

Técnicas	PORCENTAJES SEGÚN POSICIÓN								
	P0	P1	P2	P3	P4	P5	P6	P7	P8
Entrevista	5,6	77,8	5,6	11,1	--	--	--	--	--
Observación	5,6	11,1	44,4	33,3	5,6	--	--	--	--
Autoobservación	61,1	--	5,6	11,1	11,1	11,1	--	--	--
Autoinforme	22,2	11,1	22,2	16,7	27,8	--	--	--	--
Psicométricas	50,0	--	11,1	16,7	11,1	11,1	--	--	--
Objetivas	50,0	--	--	--	5,6	27,8	16,7	--	--
Subjetivas	66,7	--	--	--	5,6	5,6	11,1	11,1	--
Proyectivas	61,1	--	--	--	5,6	5,6	5,6	11,1	11,1

Pregunta: ¿Qué TÉCNICAS básicas de recogida de información cree que deberían incluirse necesariamente en la docencia de grado de la asignatura de Evaluación Psicológica? Escribalas por orden de importancia.

Nota. "P0": Porcentaje de aquellos que no escriben la técnica en ninguna de las posiciones. "P1" a "P8":

Porcentaje de personas que incluyen la técnica desde la posición 1 hasta la 8.

Cuadro 32. Test que deben incluirse en el programa docente de la asignatura Evaluación Psicológica desde la opinión de los docentes

Tests	Frecuencias Juicios					% Personas según Posición				
	Frec	%	Adm	Corr	Int	P1	P2	P3	P4	P5
Wechsler	21	13,5	90,5	90,5	47,6	83,3	16,7	--	11,1	--
MMPI-2	15	9,7	93,3	86,7	40,0	--	16,7	16,7	5,6	16,7
16PF-5	11	7,1	100	81,8	63,6	--	5,6	11,1	5,6	16,7
MCMII-III	9	5,8	88,9	88,9	33,3	--	--	5,6	11,1	5,6
PMA	8	5,2	100	100	75,0	--	11,1	5,6	11,1	--
BFQ	8	5,2	87,5	75,0	87,5	--	5,6	5,6	11,1	--
Raven	7	4,5	100	100	100	--	11,1	16,7	5,6	--
MSCA	6	3,9	100	83,3	66,7	--	11,1	5,6	--	--
DAT-5	5	3,2	100	80,0	80,0	5,6	5,6	5,6	--	5,6
EPQ-R	5	3,2	100	80,0	80,0	--	--	--	11,1	5,6
Kuder-C	5	3,2	100	80,0	80,0	--	--	--	5,6	--
Neo-PI-R	4	2,6	100	75,0	25,0	5,6	--	--	--	--
Factor G	3	1,9	66,7	66,7	66,7	--	5,6	--	--	5,6
SCL-90-R	3	1,9	100	100	100	--	5,6	--	--	--
ITPA	3	1,9	66,7	100	100	--	--	5,6	--	--
BDI	3	1,9	100	66,7	33,3	--	--	5,6	--	--
K-ABC	3	1,9	66,7	100	66,7	--	--	5,6	--	--
STAI	3	1,9	100	66,7	66,7	--	--	--	--	5,6
EPA-2	3	1,9	100	66,7	66,7	--	--	--	--	--
D-48/D-70	2	1,3	100	100	100	5,6	--	--	5,6	--
LURIA-DNA	2	1,3	100	100	100	--	--	--	--	5,6
CBCL	2	1,3	100	100	50,0	--	--	--	--	5,6
Bender	2	1,3	100	100	100	--	--	--	--	5,6
K-bit	1	0,6	--	--	--	--	5,6	--	--	--
Osgood	1	0,6	--	--	--	--	--	5,6	--	--
Peabody	1	0,6	--	--	--	--	--	--	5,6	--
ISRA	1	0,6	--	--	--	--	--	--	5,6	--
EPI	1	0,6	--	--	--	--	--	--	--	--
IPP-R	1	0,6	--	--	--	--	--	--	--	--
ABS-2	1	0,6	--	--	--	--	--	--	--	5,6
BATTELLE	1	0,6	--	--	--	--	--	--	--	5,6
EFAI	1	0,6	--	--	--	--	--	--	--	--
CMMS	1	0,6	--	--	--	--	--	--	--	--
BSID	1	0,6	--	--	--	--	--	--	--	--
RORSCHACH	1	0,6	--	--	--	--	--	--	--	--
TALE	1	0,6	--	--	--	--	--	--	--	--
REY	1	0,6	--	--	--	--	--	--	--	--
YELA	1	0,6	--	--	--	--	--	--	--	--
Rotter	1	0,6	--	--	--	--	--	--	--	--
Barcelona-R	1	0,6	--	--	--	--	--	--	--	--
Sociograma	1	0,6	--	--	--	--	--	--	--	--
TRVB	1	0,6	--	--	--	--	--	--	--	--
SOSIA	1	0,6	--	--	--	--	--	--	--	--
Naipes-G	1	0,6	--	--	--	--	--	--	--	--
Polígrafo	1	0,6	--	--	--	--	--	--	--	--
Total	155	100								

Pregunta: Señale a continuación los TESTS (nombre completo y autor) cuya inclusión considera esencial en la docencia de grado de la asignatura de Evaluación Psicológica, por orden de importancia.

Pregunta: Por favor, concrete para cada uno de los tests que ha escogido, el nivel de profundidad con que debería ser abordado en el grado (señale la/s opción/es que considere oportuna/s): Administración, Corrección, Interpretación.

Como se puede observar en los resultados de la encuesta, la entrevista es una técnica en relación a la cual todos los profesores enfatizan la importancia de su enseñanza con un alto nivel de acuerdo. Además, las técnicas observacionales, las objetivas, proyectivas, subjetivas, psicométricas y autoinformes son consideradas como técnicas relevantes para su inclusión en la docencia de grado del área de evaluación psicológica.

Dentro de los tests específicos resaltados en la encuesta por los docentes de evaluación de varias universidades españolas cabe destacar que 21 de ellos forman parte del programa actual de la asignatura Troncal “Evaluación Psicológica” en la Facultad de Psicología de la Universidad del País Vasco. Los instrumentos de evaluación infantil incluidos en el listado forman parte de la asignatura optativa “Psicodiagnóstico en la infancia y la adolescencia”.

Cuadro 33. Test del listado elaborado por los docentes que forman parte del programa de la asignatura “Evaluación Psicológica” en la Facultad de Psicología de la UPV

Wechsler	Escalas de Inteligencia de Wechsler (WAIS-III , WISC IV , WPPSI; D. Wechsler)
MMPI-2	Inventario Multifásico de Personalidad de Minnesota (MMPI-2; S.R. Hathaway y J.C. Mcklinley)
16PF-5	Cuestionario Factorial de Personalidad (16PF-5; R. B. Cattell, A.K.S Cattell y H.E.P. Cattell)
PMA	Aptitudes Mentales Primarias (PMA; L. L. Thurstone y Th. G. Thurstone)
BFQ	Cuestionario BIG-FIVE (BFQ; G. V. Caprara, C. Barbaranelli y L. Borgogni)
Raven	Test de Matrices Progresivas (RAVEN; J. C. Raven)
DAT-5	Tests de Aptitudes Diferenciales (DAT-5; G. K Bennet, H. G.Seashore y A. G. Wesman)
EPQ-R	Cuestionario de Personalidad de Eysenck - Revisado (EPQ-R; de H.J. Eysenck y S.B.G. Eysenck)
Neo-PI-R	Inventario de Personalidad Neo-PI-Revisado (Costa y McGrae)
SCL-90-R	Cuestionario de 90 Síntomas (SCL-90-R; L.R. Derogatis)
BDI	Beck Depression Inventory (BDI; Beck y cols.)
STAI	Cuestionario de Ansiedad Estado/Rasgo (STAI; C.D. Spielberger, R.L.Gorsuch y R.E. Lushene)
EPA-2	Evaluación del Potencial de Aprendizaje (EPA-2; R. Fernández-Ballesteros, M.D. Calero, J.M. Campllonch y J. Belchi)
D-48/D-70	Test de Dominós (D-48; P.Pichot y D-70; F. Kowrousky y P. Rennes)
K-BIT	Test Breve de Inteligencia de Kaufman (K-BIT; A. S. Kaufman y N. L. Kaufman)
Osgood	Diferencial Semántico (Osgood)
ISRA	Inventario de Situaciones y Respuestas de Ansiedad (ISRA; J.J.Miguel Tobal y A.R. Cano Vindel)
Rorschach	Psicodiagnóstico de Rorschach (H. Rorschach y H. Zulliger)
Sociograma	Sociograma (J.D. Coie, K.A. Dodge y H. Coppotelli)
Naipes-G	Naipes-G (N. García Nieto y C. Yuste)
Polígrafo	Polígrafo

7.10. HACIA UNA PLURALIDAD TÉCNICO-METODOLÓGICA EN EVALUACIÓN PSICOLÓGICA

En la misma dirección que la señalada por Forns (1993) considero que la aceptación de la complejidad de la conducta humana implica atender a sus múltiples factores potencialmente causales. Esto conduce a un tipo de análisis que pone de relieve los distintos valores o pesos causales que reciben los elementos potencialmente explicativos incluyendo acontecimientos situacionales, variables fisiológicas y genéticas, cambios en las variables... Implica también atender a posibles efectos causales interactivos, a la forma que toman las relaciones funcionales y a las relaciones bidireccionales entre las alteraciones de conducta y los factores causales.

La psicología actual tiende a aceptar que la conducta es un sistema con un entramado de variables, cuya influencia es interaccional y dinámica, por lo que la medida de tal sistema debe conllevar el análisis del conjunto de sus elementos. Las formas de análisis deben ajustarse a la búsqueda de la explicación de la conducta en forma de estructura o de organización, siendo los métodos de análisis multivariados adecuados a este objetivo. El uso de los sistemas estructurales permite teorizar acerca de la organización y efectos de unas variables sobre otras, e incorpora en los análisis variables observables y variables hipotéticas, latentes. Se trata del uso de los modelos causales como forma de explicar la organización de las probabilidades de covariación que mantienen distintas variables causales con la conducta objetivo. Si el análisis de la conducta de un sujeto es complejo, el análisis de acciones ejercidas en grupos o en contextos acentúa aún más la complejidad. Por ello, las propuestas multimétodos, pluritécnicas, multiaxiales y de análisis causales, así como las metodologías denominadas de triangulación, pueden constituirse en el marco integrativo idóneo para el trabajo pluridisciplinar de los psicólogos aplicados. Lo que es actualmente imprescindible es la precisión y la objetividad de los análisis y la búsqueda de consistencia lógica de los diagnósticos (Forns, 1993). Este debiera ser el objetivo común de las actuales escuelas psicológicas.

La tendencia a la integración conceptual para explicar la conducta tiene también importantes repercusiones en el dominio de las técnicas de medida. La respuesta al reto planteado de dar rigor a los análisis de conducta implica la adopción de una actitud abierta a usar los métodos y técnicas que se muestren más aptos para el análisis en cuestión. Este uso de múltiples técnicas plantea el tema de su referencia teórico-conceptual. En la actualidad, el uso de técnicas de evaluación vinculadas a diferentes enfoques psicológicos es bastante frecuente. Si bien en el plano teórico esto puede plantear un conflicto, el psicólogo aplicado, en una perspectiva ecléctica, en muchas ocasiones tiende a usarlas complementariamente, ahora bien, interpretando cada una en su contexto teórico pertinente, y dando respuestas desde distintos ángulos a los problemas que le son consultados.

Uno de los objetivos actuales, subyacentes al uso de pluritécnicas y plurimétodos, es el de conseguir datos informativos provenientes de diferentes métodos (observación, escalas de autoinforme, pruebas objetivas...) a fin de determinar a posteriori, en qué medida cada una de estas técnicas aporta información relevante para la identificación de una alteración. Los métodos estadísticos de regresión múltiple podrían determinar qué peso corresponde a cada tipo de técnicas de obtención de información, en función de su capacidad predictiva frente al criterio establecido. Estos análisis permitirán a los profesionales conocer la clase de técnicas más adecuadas en función de las hipótesis diagnósticas que desearan contrastar y plantear la

combinatoria (o batería) ideal de pruebas utilizables en relación con la comprobación diagnóstica de determinadas alteraciones (Forns, 1993).

No obstante, hay que señalar, que en el momento actual la Evaluación Psicológica se caracteriza por la presencia de inadecuaciones entre la conceptualización y las técnicas de análisis. El psicólogo aplicado se encuentra con grandes problemas en su praxis evaluativa ya que en ocasiones cuenta con marcos teóricos muy elaborados, y en la práctica carece de instrumentos apropiados, mientras que en otras ocasiones le sucede lo contrario.

Otro tipo de dificultad proviene de la dificultad de realizar precisos análisis interaccionales de la conducta. Actualmente se comparte la creencia de que la conducta está influida por muchos sistemas en los cuales el individuo está inmerso, sin embargo, es prácticamente imposible un análisis extensivo de todos ellos y sus mutuas interacciones. Los estudios interaccionales precisan de taxonomías que definan diferentes parámetros ambientales y requieren una metodología compleja de tipo multivariante. Por ello, es difícil realizar este tipo de análisis en el campo de la evaluación individual, quedando restringido a nivel de investigación y con diseños grupales. Las dificultades metodológicas para la realización del análisis interaccional ha promovido la aceptación de metodologías cualitativas y descriptivas en la valoración de programas.

Por otra parte, la psicología actual está más preocupada por definir las conductas en términos de *objetivos de logro* más que en términos de *normas*. La evaluación criterial determina en qué medida un sujeto se acerca a los estándares criterios fijados para él o para un grupo concreto, de tal modo que el avance o retroceso personal en relación a tales criterios pueda ser interpretado directamente. La evaluación referida a normas compara las actuaciones de un sujeto en relación a un grupo, cuya eficiencia se constituye en modelo distribucional; la actuación de un individuo no tiene nunca valor *per se*, sino en tanto se refiere a las normas del grupo de referencia. En estos parámetros el evaluador se encuentra con un problema: la escasez de pruebas referidas a criterios y la dificultad de construir escalas criterios para cada uno de los casos particulares.

Si bien la Evaluación Psicológica ha asumido los grandes cambios conceptuales influidos por la Psicología Cognitiva y por el Análisis de Acciones Psicosociales, sin embargo, requiere de renovaciones tecnológicas adecuadas a los actuales objetivos de estudio y medida.

En relación al empleo de distintos procedimientos diagnósticos utilizados en la práctica psicológica, en un primer momento se cuestionaron las técnicas proyectivas en favor de los tests psicológicos. Después, con la revolución conductista se promocionaron determinados procedimientos, técnicas, o estrategias de evaluación conductual en detrimento de los tests psicológicos. Posteriormente, se ha procedido a una integración que tiene que ver con la necesidad de utilizar procedimientos rigurosos y diferencialmente validados para cada condición diagnóstica. Sin embargo, cuando se analizan los procedimientos más utilizados en la práctica diagnóstica, esta realidad no se constata. Estudios recientes muestran un gran predominio de la entrevista y un mayor uso de procedimientos usualmente no estandarizados (escalas de calificación de elaboración propia). Los tests psicológicos se utilizan con mucha frecuencia pero su nivel de utilización depende del ámbito de la práctica. Otros procedimientos muestran una tendencia similar, aunque su uso es menos frecuente (observación sistemática, registros psicofisiológicos, juego de roles...).

La evaluación psicológica ha estado provista de variados y múltiples instrumentos de evaluación que han dado satisfactorias respuestas a las necesidades de los profesionales en

distintos ámbitos. No obstante, se hace necesario abrir o, en palabras de Trull (2007), expandir el concepto de evaluación psicológica y sus métodos y estrategias de trabajo a tecnologías y métodos innovadores que complementen las necesidades actuales en evaluación. Para este autor, cuya revisión de las nuevas tecnologías aplicadas al ámbito de la evaluación merece tenerse en cuenta, plantea entre otros por su especial relevancia: el uso de la Entrevista computerizada (Garb, 2007), Diarios electrónicos inicialmente utilizados en el contexto ecológico y en la actualidad en la evaluación clínica (Piasecki, Hufford, Solhan y Trull, 2007), el uso de Biosensores en la evaluación del comportamiento fisiológico y de funciones biológicas (Haynes y Yoshioka, 2007), así como otras técnicas de neuroimagen demostradamente válidas para los profesionales de la salud (Miller, Elbert, Sutton, y Heller, 2007)

7.11. RESUMEN. TÉCNICAS DE RECOGIDA DE INFORMACIÓN EN EVALUACIÓN PSICOLÓGICA

A modo de síntesis del capítulo a continuación se clarifican las diferencias entre técnica y test, así como algunas taxonomías propuestas para su clasificación (Cuadro 34). Posteriormente, se sintetizan algunas ideas básicas referentes a la clasificación que proponemos subrayando las características, ventajas y dificultades de la entrevista (Cuadro 35), los autoinformes (Cuadro 36), las técnicas observacionales (Cuadro 37), las técnicas proyectivas (Cuadro 38), las técnicas subjetivas (Cuadro 39), las técnicas objetivas (Cuadro 40), y las técnicas psicométricas (Cuadro 41).

Cuadro 34. Las técnicas en evaluación psicológica y sus taxonomías

LAS TÉCNICAS EN EVALUACIÓN PSICOLÓGICA Y SUS TAXONOMÍAS	
Técnica - Test	<i>Técnica-Instrumento:</i> • Procedimiento de recogida de información no cuantificado ni tipificado (Pelechano, 1976). • Dispositivo de recogida de información sin que necesariamente requiera tipificación o permita comparaciones intersujetos (Fernández-Ballesteros, 2004). <i>Test:</i> • Instrumento sistemático y tipificado que compara la conducta de dos o más personas (Pelechano, 1976). • Procedimiento de recogida de información en el que la administración, corrección y valoración son estándar, y del que se derivan puntuaciones normativas (Del Rincón y otros, 1995; Fernández-Ballesteros, 2004). • Procedimiento sistemático para obtener muestras de conducta significativas del funcionamiento cognitivo o afectivo de la persona y para puntuar y evaluar estas muestras de acuerdo a normas tipificadas (Urbina, 2007)
Taxonomías	<i>Taxonomía Pervin (1979)</i> • Tests proyectivos • Técnicas subjetivas • Tests psicométricos • Tests objetivos

	<i>Taxonomía Fernández-Ballesteros (1992)</i> • Técnicas de observación • Técnicas de autoinforme • Técnicas de entrevista • Técnicas subjetivas • Técnicas proyectivas • Técnicas objetivas <i>Taxonomía Caballo (2005)</i> • La entrevista • Cuestionarios, inventarios, escalas • La observación conductual • La evaluación cognitiva • El autorregistro • La evaluación por los demás • Las medidas análogas (de lápiz y papel, de audio, de video, de actuación, de representación de papeles) • La evaluación psicofisiológica <i>Taxonomía Calero (2007)</i> • Técnicas objetivas - Observación • Técnicas objetivas - Estructuradas: técnicas objetivas, registros psicofisiológicos, tests psicométricos de ejecución. • Técnicas voluntarias estructuradas: autoobservación, autoinformes psicométricos • Técnicas voluntarias - los autoinformes: entrevista. • Técnicas voluntarias no-estructuradas: técnicas subjetivas • Técnicas objetivas no-estructuradas: técnicas proyectivas

Cuadro 35. La entrevista psicológica

LA ENTREVISTA PSICOLÓGICA	
Definición	Una relación interpersonal entre dos o más personas; una vía de comunicación simbólica bidireccional, preferentemente oral; con unos objetivos prefijados y conocidos, y una asignación asimétrica de los roles entrevistador-entrevistado. Es un procedimiento ampliamente utilizado con distintos objetivos: conseguir información u ofrecer información en la evaluación e intervención psicológica; indagar sobre características de personalidad, estrategias, incidentes críticos, habilidades, competencias, y datos biográficos en procesos de selección; indagar sobre los aspectos motivacionales, intereses y competencias en los procesos de orientación; investigar el conocimiento que pone en juego un experto cuando resuelve problemas de su competencia; orientar o asesorar y comunicar a clientes o pacientes los resultados de cualquier proceso evaluativo. Es un instrumento destinado a recoger datos mediante el autoinforme o heteroinforme del sujeto evaluado.

Clasificación	• En función de los objetivos: a) Diagnósticas, Terapéuticas, De consejo u orientación; De investigación; b) Entrevistas de recogida de información y entrevistas de devolución de la información. • En función del grado de estructuración: Libre, Semiestructurada; Estructurada. • En función del marco teórico: Psicodinámico, Fenomenológico, Conductual.
Habilidades del Entrevistador	Habilidades • Habilidades empáticas • Habilidades no verbales • Conocerse a uno mismo • Habilidades para preguntar • Operacionalizar las informaciones • Parafrasear • Reflejar sentimientos • Dirigir la entrevista • Mantener la interacción • Resumir Actitudes relacionadas con las buenas habilidades para la entrevista: • Cordialidad • Autenticidad • Aceptación • Comprensión • Franqueza • Honestidad • Imparcialidad
Características psicométricas	• Validez de contenido: Se incrementa en las entrevistas estructuradas. • Validez de criterio: Se incrementa en las entrevistas con criterios objetivos. • Fiabilidad interevaluadores: Coeficientes más altos cuanto se valoran aspectos conductuales, objetivos y observables, y más bajos cuando se valoran impresiones subjetivas e interpretaciones personales. Relación directa con el grado de entrenamiento y experiencia de los evaluadores. • Fiabilidad test-retest: Mayor concordancia en las apreciaciones negativas que en las positivas.
Fuentes de distorsión	• La habilidad del entrevistador. • Los estatus de ambos. • Los efectos de reactividad. • Los contenidos sobre los que se van a tratar. • El desconocimiento o la falta de información...
Estrategias facilitadoras	<i>Sugerencias:</i> Tener un plan definido, conocer las variables de criterio, organizar la entrevista adecuadamente en relación al tiempo disponible; facilitar la motivación del entrevistado,

	garantizar la privacidad de la información, dejar hablar al entrevistado, ajustar el lenguaje al lenguaje del entrevistado, tener conciencia de los propios prejuicios para evitar su influencia, registrar adecuadamente los datos obtenidos. <i>Asegura una mayor objetividad:</i> La estructuración de preguntas y respuestas. El uso de categorías de registro definidas. La ordenación de las respuestas en esquemas o pautas. La unidad de perspectivas de juicio entre los distintos entrevistadores. El entrenamiento en el registro y la tabulación de la información. <i>Asegura una mayor fiabilidad:</i> La estimación de las bases motivacionales de la respuesta. La utilización de racimos de preguntas, más que de preguntas aisladas. La clara especificación temporal de las preguntas. Preguntar por acontecimientos objetivos o indicadores observables de la conducta. <i>Asegura una mayor validez:</i> Dar instrucciones y aclaraciones que sitúen al entrevistado en los diversos temas. La transparencia en las preguntas y la utilización de un lenguaje comprensible. La preferencia de indicadores observables de la conducta. Los intentos de detectar y neutralizar tendencias de respuesta. Asegurar la confidencialidad. Abrir las posibilidades de contrastación de las respuestas a través de otras fuentes de información. <i>Recomendaciones para manejar entrevistas interétnicas, interculturales y entre clases:</i> • Incrementar la conciencia cultural (ser sensible a las diferencias culturales, de clase y étnicas) • Conocerse a uno mismo (estereotipos, prejuicios...) • Ser flexible (evitar ideas preconcebidas, estar abierto a aceptar una perspectiva diferente a la suya) • Ver más allá de uno mismo (ser capaz de apreciar la perspectiva del otro)

Cuadro 36. Los autoinformes

LOS AUTOINFORMES	
Definición	Un mensaje que el sujeto emite sobre cualquier tipo de manifestación propia (conductas motoras, pensamientos, percepción de una situación, experiencias...).
Clases de variables que evalúan	• Rasgos o dimensiones de personalidad. • Estados. • Repertorios y procesos cognitivos. • Repertorios clínicos conductuales.
Tipos de autoinformes	• Entrevista, Autobiografía, Historia Clínica. • Cuestionarios, Inventarios, Escalas (no estandarizados y estandarizados-psicométricos)

	• Autorregistros. • Pensamientos en voz alta.
Fuentes de error	• Del evaluado: Simulación o engaño deliberado, Deseabilidad social, Aquiescencia, Tendencia de respuesta, Motivación, Dificultad para la introspección, Dificultad de comprensión de los ítems... • Del evaluador: Características personales (sexo, aspecto, raza, estilo personal...). • Del instrumento: Validez, Fiabilidad. • De la situación: Existen autoinformes más situacionales mientras que otros son más independientes del momento presente.

Cuadro 37. Las técnicas observacionales

LA OBSERVACIÓN	
Definición	• La observación es la estrategia fundamental del método científico, y, por tanto, cualquier técnica de recogida de información en evaluación psicológica conlleva, de una u otra forma, observación. • Es observación cualquiera de los procedimientos y técnicas de recogida de información que requieren la percepción deliberada de comportamientos emitidos por un/os sujeto/s (a distintos niveles de complejidad e inferencia), realizada por un observador entrenado (participante o no participante) mediante protocolos preparados al efecto (de mayor o menor estructuración) que permitan una anotación sistemática, en una situación natural o artificial.
Objetivo	• A través de la observación se pretende recoger mediante un registro objetivo, sistemático y específico, las conductas que se producen en un determinado contexto, sometiendo dicha información a una adecuada codificación y análisis.
Unidades de observación	• Continuo del comportamiento. • Atributos o rasgos de personalidad. • Conductas. • Interacciones. • Productos de conducta.
Técnicas de Registro	• Registro narrativo. • Escalas de apreciación. • Catálogos de conducta: Registros de conducta y Matrices de interacción. • Códigos o sistemas de categorías. • Registro de productos de conducta. • Procedimientos automáticos de registro.
Tipos de muestreo	• Muestreo ad libitum (experiencias no sistematizadas, notas de campo). • Muestreo de eventos (selección como unidades de la muestra de todas las ocurrencias de una conducta o gama de ellas que tengan lugar a lo largo de las sesiones establecidas).

	• Muestreo focal (se registran todas las acciones de un sujeto previamente seleccionado, y las que se le dirigen). • Muestreo de tiempo (muestreo instantáneo o de puntos de tiempo, y muestreo de intervalos).
Fuentes de error o sesgos	• Del sujeto observado: reactividad, autoreactividad, características del sujeto... • Del observador: grado de participación, expectativas, características personales, reactividad recíproca, entrenamiento. • Del sistema de observación: tipo de registro, tipo de muestreo, procedimientos eléctricos o mecánicos de registro de los datos...
Planificación de la Observación	• ¿Requiere el caso observación? ¿Propósito de la observación? • ¿Qué unidades de observación están implícitas en la hipótesis sobre el caso? (continuo de comportamiento, atributos, conductas, interacciones, productos de conducta...). ¿Características y tipo de variable objeto de observación? • ¿Qué unidades de medida van a ser utilizadas? (frecuencia, ocurrencia, duración...). • ¿Mediante qué técnicas van a ser recogidas las observaciones? (registros narrativos, escalas de apreciación, catálogos de conducta...) ¿Calidad de los datos? • ¿Quién o quiénes van a ser observados? • ¿Dónde se va a observar? (situaciones naturales o artificiales). • ¿Durante cuánto tiempo se realizará la observación? (tiempo total de observación, número de sesiones, duración de las sesiones, periodos de habituación...). • ¿Quién o quiénes van a observar? (procedimiento de formación de los observadores). • ¿Medios disponibles? • ¿Cómo van a apreciarse la bondad de las observaciones? (fiabilidad y validez).
Fases lógicas del procedimiento en metodología observacional	• Fase exploratoria. • Disposición previa (mantenimiento de la constancia intersesional, mantenimiento de la constancia intrasesional, tratamiento de las disrupciones temporales, temporalización e identificación de la sesión de observación). • Plan de muestreo observacional. • Elaboración del instrumento (sistema de categorías, formatos de campo, combinación de formatos de campo y sistemas de categorías, escalas de estimación o apreciación). • Registro y codificación. • Métrica del registro u obtención de parámetros. • Control de calidad del dato. • Análisis de datos.

Cuadro 38. Las técnicas proyectivas

LAS TÉCNICAS PROYECTIVAS	
Definición	- Instrumentos sensibles para revelar aspectos inconscientes de la persona; son multidimensionales, enmascaradas, presentan un material ambiguo y requieren respuestas de la imaginación. - Medidas que requieren una respuesta no estructurada en orden a efectuar una evaluación de la personalidad o cognición del sujeto examinado.
Hipótesis del concepto de proyección	- La hipótesis proyectiva sostiene que un individuo da estructura a estímulos que no están estructurados de una manera consistente con el propio patrón único de necesidades, temores, deseos, impulsos, conflictos y formas de percibir y responder conscientes e inconscientes. De acuerdo con la hipótesis proyectiva, las interpretaciones de un estímulo ambiguo o vago reflejan las necesidades propias de un sujeto, sus sentimientos experiencias, condicionamiento previo y procesos de pensamiento entre otros. - La revisión del concepto de proyección ha evidenciado distintas hipótesis del concepto destacando varias hipótesis desde las que se entiende este concepto de distinto modo: Hipótesis proyectiva clásica. Hipótesis holística-funcional. Hipótesis de la desviación. Hipótesis del nivel de adaptación. Hipótesis de los estados directivos. Hipótesis de la actividad perceptiva. Hipótesis de la situación experimental. Hipótesis cognitivo-experimental. Hipótesis cognitiva-aplicada. Hipótesis Existencial-Humanista.
Clasificación	<i>Técnicas estructurales</i>: Se presenta material visual de escasa estructuración que el sujeto debe estructurar diciendo lo que ve, o a lo que se puede parecer (Rorschach). <i>Técnicas temáticas</i>: Se presenta material visual con distintos grados de estructuración formal de contenido humano o parahumano sobre los cuales el sujeto debe narrar una historia (Test de Apercepción Temática de Murray, Test de Relaciones Objetales de Phillipson). <i>Técnicas expresivas</i>: Se da al sujeto la consigna verbal o escrita de dibujar una/s figura/s (por ejemplo, el Test de la Figura Humana, el Test del Árbol, el Test de la Familia...). <i>Técnicas constructivas</i>: Se aporta material concreto que el sujeto debe de organizar en base a distintas consignas (por ejemplo, test del pueblo). <i>Técnicas asociativas</i>: Se aporta una consigna verbal o escrita por la que el sujeto ha de manifestar verbalmente sus asociaciones frente a palabras, frases o cuentos (por ejemplo, las Fábulas de Duss, o el Test de las Frases Incompletas).
Tres vías de desarrollo	- Reactualización desde la perspectiva conductual cognitiva. - Reactualización desde la psicología de la percepción. - Reactualización desde la psicometría.

Cuadro 39. Las técnicas subjetivas

TÉCNICAS SUBJETIVAS	
Definición	• Una serie de métodos orientados a recoger y organizar información acerca de la estructura psicológica, los contenidos y los procesos de las visiones subjetivas o significados personales acerca de uno mismo y el mundo. • En contraste con los métodos que tienen como objeto clasificar al sujeto en unas determinadas dimensiones (por ejemplo extraversión-introversión) o categorías (trastorno mental) ya establecidas de antemano, las técnicas subjetivas pretenden captar sus construcciones personales, en definitiva, estudiar sistemáticamente la estructura de su subjetividad.
Características	• La fuente de datos puede ser el propio sujeto que se califica o clasifica a sí mismo o a objetos, personas o conceptos significativos, o también personas allegadas al sujeto en exploración que realizan esa misma tarea pero referida a éste. • El material suele ser semiestructurado y no enmascarado. No suele estar tipificado, es flexible y se adapta a los objetivos que se persiguen en cada caso. • La respuesta del sujeto es voluntaria. • El tipo de tarea requerida es la de descripción, calificación o clasificación del propio sujeto o de otros objetos o abstracciones. • La manipulación de los resultados no está reglada, pudiendo optar por planteamientos cualitativos o cuantitativos en la elaboración de respuestas. • Construidas desde planteamientos teóricos fenomenológicos y/o cognitivos, en la actualidad se emplean desde perspectivas teóricas muy distintas (conductismo, psicoanálisis...).
Técnicas representativas	• La técnica de clasificación Q de Stephenson (1953). • La técnica del diferencial semántico de Osgood (1952). • El test de los constructos personales de Kelly (1955). • Las listas de adjetivos. • El método biográfico.

Cuadro 40. Las técnicas objetivas

TÉCNICAS OBJETIVAS	
Definición	• Procedimientos de recogida de información de eventos psicológicos observables o amplificables, que en gran parte de los casos no son controlables (al menos sin entrenamiento), los cuales se aplican mediante sofisticados aparatos que permiten una administración, registro, puntuación y análisis objetivo (generalmente mecánico o electrónico) sin la intervención del evaluador. • Características: ➢ Requieren una instrumentación y un material estándar cuya aplicación se realiza en condiciones estructuradas y de máximo control (objetividad en el procedimiento y artificialidad). ➢ El sujeto no puede modificar sus respuestas de forma voluntaria.

	➤ Las respuestas del sujeto pueden ser registradas, codificadas y procesadas sin que medie la opinión o criterio del evaluador.
Clasificación	<i>Instrumentación Perceptiva</i> • Estesiómetro • Test de discriminación de dos puntos • Dolorímetro electrónico portátil • Medida de la Percepción de la Profundidad. • Percepción de la verticalidad: El test del Marco y la Varilla. <i>Instrumentación Cognitiva</i> • Polirreactígrafo. • Taquistoscopio. • Tambores de memoria... <i>Instrumentación Motora</i> • Omega. • Aparato de precisión manual de Purdue. • Octobinexímetro. • Simulador de conducción. • Silla vibrátil... <i>Instrumentación Psicofisiológica</i> • Polígrafo. • Sistemas modulares computerizados. • Aparatos de registro portátiles...
Características psicométricas	• Validez interna: Apreciable. • Validez externa, ecológica: Dificultades.

Cuadro 41. Las técnicas psicométricas

TÉCNICAS PSICOMÉTRICAS	
Definición	Las técnicas psicométricas o tests psicológicos tienen como características fundamentales: • <i>Estandarización</i>: El material, instrucciones, procedimiento de aplicación e interpretación están estandarizados, es decir, son idénticos para todos los sujetos evaluados. • <i>Corrección objetiva</i>: El resultado obtenido por el sujeto es independiente del profesional que realiza la evaluación, ya que el procedimiento de corrección está también estandarizado y no está sesgado por el conocimiento del evaluador. • <i>Resultado cuantitativo</i>: Los resultados ofrecen un resultado numérico. • <i>Tipificación</i>: La interpretación sólo tiene sentido a partir de la comparación de las

	puntuaciones obtenidas por el sujeto con un grupo de referencia normativo, es decir, con un grupo que se establece como norma.
Clasificación	Test de inteligencia: • Tests individuales. • Test colectivos: inteligencia general frente a batería de aptitudes. Test de personalidad: • Pruebas multirasgo. • Pruebas monorasgo.
Presente y futuro de los test psicológicos	Aunque en diversos momentos históricos se han criticado, los tests psicológicos siguen siendo hoy uno de los instrumentos más utilizados. Algunas de las razones que pueden explicar este hecho son: • Uno de los objetivos de la evaluación psicológica es la selección profesional y/o educativa. Para este fin, es necesario utilizar instrumentos que posibiliten: a) la evaluación de gran número de sujetos, al mismo tiempo, en las mismas condiciones para todos ellos; y b) la comparación. Estos requisitos son los que guían la construcción de tests psicométricos colectivos, y por tanto, en este campo siguen siendo instrumentos de necesaria aplicación. • Cuando el objetivo de la evaluación psicológica es el diagnóstico, en muchas ocasiones se requiere contrastar correlacionalmente determinadas hipótesis antes de pasar a una evaluación más precisa del caso. Los tests psicológicos basan su interpretación en criterios normativos y por ello pueden ser útiles para este tipo de comprobación. • En un proceso de orientación escolar y/o laboral se puede solicitar al evaluador una descripción del individuo, y ello implica establecer perfiles de aptitudes y personalidad que ayuden a tomar decisiones. Muchos tests psicológicos en los últimos años se han orientado a la realización de perfiles comparativos. Por consiguiente, el uso de tests psicológicos es adecuado para diversos objetivos de la evaluación, especialmente cuando se requiere la comparación del individuo con un grupo de referencia normativo. Por ello, se puede predecir que el uso de los tests psicológicos se perpetuará en el futuro, siempre que cumplan con las adecuadas garantías de fiabilidad y validez, y se construyan utilizando los nuevos desarrollos científicos (aplicación de los modelos de procesamiento a los tests de inteligencia y aptitudes, aplicación de la TRI, utilización de test adaptativos informatizados, uso del ordenador con el positivo efecto que está produciendo en esta metodología...)

CAPÍTULO 8.

LINEAS DE FUTURO: EVALUACIÓN INFORMATIZADA Y EVALUACIÓN DE PROGRAMAS

8.1. Introducción

8.2. La evaluación psicológica informatizada

8.2.1. Aplicaciones para la administración de pruebas

8.2.1.1. Aplicación informatizada de instrumentos tradicionales

8.2.1.2. Aplicación de instrumentos para ser administrados únicamente por ordenador

8.2.1.3. Aplicaciones combinando ordenador y vídeo

8.2.1.4. Tests adaptativos informatizados (tests a medida)

8.2.1.5. Instrumento auxiliar en la aplicación de técnicas de potencial de aprendizaje

8.2.1.6. Desarrollo de programas informáticos integrales

8.2.1.7. Aplicación de pruebas basadas en criterios (pruebas de competencia mínima)

8.2.2. Aplicaciones para la corrección - puntuación de pruebas

8.2.3. Aplicaciones para la interpretación de pruebas

8.2.3.1. Tipos de informes computerizados

8.2.3.2. Los sistemas expertos

8.2.3.3. Programas informáticos en el mercado español

8.2.4. Impacto y polémica de la evaluación informatizada

8.2.4.1. La administración informatizada

8.2.4.2. La interpretación informatizada

8.2.4.3. Limitaciones actuales de la evaluación informatizada

8.3. La evaluación de programas, intervenciones o tratamientos

8.3.1. Bases conceptuales

8.3.1.1. Definición y objeto de estudio

8.3.1.2. Criterios para la evaluación de programas

8.3.1.3. Modelos en la evaluación de programas: evaluación sumativa y formativa

8.3.1.4. Tareas y fases del proceso en la evaluación de programas

8.3.2. Cuestiones metodológicas

8.3.2.1. Principales diseños en la investigación evaluativa

8.3.2.2. Procedimientos e instrumentos para evaluar el cambio

8.3.2.3. Problemas metodológicos relacionados con el programa

8.3.2.4. El análisis de los datos en la evaluación de los efectos del programa

8.3.3. Ámbitos de aplicación de la evaluación de programas en España

8.4. Algunas predicciones futuras sobre la evaluación psicológica y el rol del psicólogo en España

8.5. Resumen: evaluación informatizada, evaluación de programas y predicciones de futuro

8.1. INTRODUCCIÓN

En este capítulo se describen dos líneas de futuro, dos líneas en expansión en evaluación psicológica, la evaluación informatizada y la evaluación de programas. En primer lugar, se analiza la situación de creciente utilización del ordenador en la evaluación, sus posibilidades actuales, los resultados de diversos estudios que analizan el impacto de la evaluación computerizada, así como las posibilidades de futuro que se abren con el desarrollo de nuevas tecnologías. En segundo lugar, se reflexiona sobre una reciente línea de trabajo en el campo de la evaluación psicológica, la evaluación de programas o de intervenciones preventivas, para fomentar el desarrollo (empatía, prosocialidad, creatividad...) y terapéuticas (mujeres maltratadas, tratamiento a maltratadores, intervenciones en trastornos mentales...). Esta línea ha permitido una mayor incorporación del psicólogo evaluador en distintos ámbitos profesionales desde los que en la actualidad se responde a concretas demandas y necesidades sociales (prevención de drogas, protección de menores, intervenciones psicoeducativas para la prevención de la violencia, de ayuda a inmigrantes...).

8.2. LA EVALUACIÓN PSICOLÓGICA INFORMATIZADA

En la actualidad los ordenadores se han extendido a múltiples campos y aspectos de la vida cotidiana, y este impacto se ha visto también reflejado en la Evaluación Psicológica. El empleo del ordenador ha sido debatido ampliamente, manifestándose posturas contrapuestas entre los psicólogos. No obstante, y pese a las críticas, estas herramientas, aunque en el futuro irán perfeccionándose, actualmente pueden ser una ayuda a la hora de realizar el proceso que implica la Evaluación Psicológica.

La incorporación del ordenador ha sido progresiva desde la década de los 60, pudiéndose distinguir tres fases en este desarrollo:

1. *Década de los 60. Los grandes ordenadores y los servicios de corrección e interpretación:* Hacia 1960 comenzaron a desarrollarse los primeros programas que permitían la corrección e interpretación mecanizada de las pruebas psicológicas más utilizadas. Estas aplicaciones pusieron de manifiesto las ventajas de los ordenadores, pero básicamente eran utilizados por grandes empresas especializadas en servicios de corrección-interpretación.
2. *Década de los 70. Los miniordenadores y las técnicas de automatización de conductas relevantes para intervención:* La reducción del tamaño de los ordenadores, y el abaratamiento de su coste, estimula su progresiva implantación en los servicios de salud mental americanos. No obstante, los psicólogos clínicos plantean variadas críticas que conducen progresivamente a la fabricación de programas con la finalidad de ayudar a los clínicos en el proceso de toma de decisiones, desarrollándose programas más sofisticados y complejos, los cuales permiten la administración, corrección e interpretación automatizada de una amplia batería de pruebas psicológicas sobre cuyos datos se tomarían decisiones relevantes de cara a la intervención. En este período el coste de la utilización del ordenador resultaba muy caro lo que impidió que se extendiera a los profesionales que ejercían su trabajo de forma autónoma.

3. *Desde la década de los 80. Los microordenadores y el desarrollo de la comercialización del software:* Progresivamente ha sido posible adquirir microordenadores potentes y a un bajo coste, lo que ha potenciado un gran desarrollo de software destinado a ayudar en todas las fases de proceso de evaluación psicológica.

A partir de los años cincuenta, al principio muy lentamente, en los últimos veinte años del siglo XX de forma acelerada y en la actualidad con una rapidez vertiginosa, los ordenadores han tomado parte en nuestra vida, tanto profesional como personal. En las revisiones sobre fuentes documentales, los datos referidos al uso de ordenador bien en CD-ROM bien vía Internet ocupan un lugar fundamental. La incorporación del ordenador al quehacer psicológico, no sólo como instrumento de apoyo a la investigación (en modelos de procesamiento de la información) o de análisis complejos de datos -como en las aproximaciones psicométricas- sino también en la psicología aplicada, de las organizaciones, educativa y clínica, es algo incuestionable. Es interesante ver cómo desde hace algunos años las revisiones sobre el tema abarcan incluso las psicoterapias conductual, cognitiva y dinámica (Calero y Fernández-Ballesteros, 2004; Calero 2007).

Hoy en día, los microordenadores son utilizados para administrar, puntuar, y/o interpretar muchas de las pruebas psicológicas empleadas en evaluación, en cualquiera de sus ámbitos de aplicación. Los programas de software son una incalculable ayuda puesto que realizan de forma rápida, fiable y con una gran exactitud todas las tareas relacionadas con la corrección e interpretación de tests, tareas tediosas, que conllevan gran cantidad de tiempo y en las que era fácil equivocarse. Además, los sistemas expertos proporcionan ayuda en la tarea de toma de decisiones diagnósticas.

Un indicador del progresivo impacto del ordenador en la evaluación psicológica, es el hecho de que en los últimos años ha habido un aumento espectacular en las fuentes de información sobre las aplicaciones clínicas de la informática. En los más recientes congresos del área han sido frecuentes las demostraciones sobre material de software desarrollado. Además, cada día se publican nuevos libros sobre aplicaciones de software relacionadas con la práctica clínica y escolar. Existen varias revistas dedicadas específicamente a las aplicaciones informáticas (*Computers in Human Behavior*; *Behavior Research Methods, Instruments & Computers...*), y otras muchas con secciones habituales y artículos sobre estos desarrollos (*European Journal of Psychological Assessment*; *Behavioral Research and Therapy*; *Clinical Psychology Review...*).

Otro indicador es que los manuales actuales de tests psicológicos incluyen capítulos dedicados a la evaluación psicológica asistida por ordenador (Cohen y Swerdlik, 2001; Garaigordobil, 1998; Graham y Naglieri, 2003; Kaplan y Sacuzzo, 2006; Cordero, 1997; Westmeyer, 2001), y denominaciones como "evaluación asistida por ordenador" "tests adaptativos informatizados", "diagnóstico informatizado" "computadoras y pruebas psicológicas" son cada vez de uso más frecuente y hacen referencia a una estrategia alternativa a la evaluación tradicional de variables psicológicas relevantes, determinada por: 1) el uso del ordenador como soporte, procedimiento de control y/o guía, 2) que no se limita a la evaluación cualitativa, y 3) que no sustituye totalmente al psicólogo evaluador. Este esquema responde a cualquier objetivo de evaluación informatizada, desde la aplicación de una técnica hasta la realización de un sistema experto.

Actualmente, existen una gran cantidad de aplicaciones que desde hace pocos años está teniendo la informática en la Evaluación. En relación a estas aplicaciones, la APA (American Psychological Association) publicó hace 20 años un directorio de software (Stoloff y Couch, 1989) en el que ya se citaban 354 programas de ordenador dirigidos tanto a la aplicación como a la puntuación e interpretación de técnicas de todo tipo, desde proyectivas y subjetivas hasta perceptivas y psicofisiológicas, pasando por toda una gama de posibilidades de Historia Clínica y Autoinformes (Fernández-Ballesteros y Calero, 1992).

También en España, en los catálogos de distribución de material de evaluación podemos observar cómo cada vez se dejan más de lado los sistemas de corrección tradicionales y aparecen los soportes informáticos de corrección, de aplicación (vía Internet) y de interpretación. En concreto, en el catálogo de TEA, 2009, aparecen 89 pruebas en las que se puede realizar una corrección mecanizada por ordenador, 43 permiten realizar aplicaciones en línea, 15 incorporan disquette para la corrección, 29 se pueden aplicar por Internet, 16 instrumentos informatizados, y varios programas informáticos integrales por ejemplo el TEA-System 4.0 que permite la corrección e interpretación computerizada. En relación a las técnicas objetivas, existen hoy día diversos programas informáticos dirigidos a la evaluación de diferentes variables tanto perceptivas como motoras y/o fisiológicas.

El ordenador puede usarse para las pruebas de dos formas básicas: 1) para administrar, calificar e interpretar pruebas tradicionales; y 2) para crear nuevas tareas y medir capacidades que los procedimientos tradicionales no pueden explorar (Kaplan y Saccuzzo, 2006). La incorporación del ordenador como instrumentación auxiliar en Evaluación psicológica tiene lugar en relación a tres momentos del proceso evaluador: En la administración, en la corrección-puntuación, y en la interpretación de los resultados.

8.2.1. APLICACIONES PARA LA ADMINISTRACIÓN DE PRUEBAS

En la actualidad se pueden destacar las siguientes aplicaciones en la administración de pruebas por ordenador:

8.2.1.1. Aplicación informatizada de instrumentos tradicionales

Con respecto a las versiones informatizadas de tests clásicos, podemos señalar que se han diseñado programas para aplicar directa o indirectamente cualquier tipo de pruebas de personalidad, clínicas y/o de ejecución.

- Entrevistas e Historias Clínicas: Existen numerosos programas distintos para la realización de Historias Clínicas. Pueden ser cumplimentadas por el psicólogo, por los padres del sujeto, o por el sujeto mismo.
- Cuestionarios e Inventarios Conductuales: Dirigidos a la evaluación de diversos problemas como ansiedad, estrés, depresión, dolor crónico, alcoholismo, abuso de fármacos, problemas sexuales...
- Técnicas Subjetivas: Listas de adjetivos y Técnicas de Rejilla.
- Tests de Personalidad e Intereses (MMPI, 16 PF, Strong, EPI....).
- Baterías de Aptitudes (PMA), Inventarios de conducta adaptativa, Tests de rendimiento, Tests de inteligencia...

En los últimos veinte años progresivamente han ido apareciendo más procedimientos para informatizar instrumentos tradicionales. Ejemplos de estos son el MACI (Calero y Araque, 1990a) que es una matriz para informatizar cuestionarios. En nuestro país existe un programa (Cuestionarios Individuales Informatizados) que permite informatizar en esta forma cualquier test verbal de papel y lápiz, tanto autoinforme como test de ejecución, permitiendo manipular tiempos de presentación y opciones de respuesta, y suministra información de aciertos, errores, corrección, tiempos de los sujetos evaluados (Calero y Araque, 1991a, 1992). En este contexto cabe mencionar los cuestionarios de personalidad por ordenador “visual Basic 6.0 para Windows” (Pelechano y Darías, 2000), este programa permite la corrección automática de siete cuestionarios de personalidad de V. Pelechano (EN, ENR, HE-A, LUCAD, MA, MOTEX y R3). Los programas incluyen la presentación de instrucciones y los ítems por pantalla para cumplimentación individual, tiempo de cumplimentación, cambio de respuesta por parte del sujeto, puntuación individual directa y tipificada.

En la mayoría de estas aplicaciones, el formato de presentación es similar al formato de lápiz y papel de esas técnicas. No obstante, la consideración de los tiempos de respuesta, la adaptación al sujeto de instrucciones y ayudas y de la propia prueba mediante sistemas de ramificación de ítems son algunas de las múltiples posibilidades que el ordenador ofrece (Butcher, 2003).

8.2.1.2. Aplicación de instrumentos para ser administrados únicamente por ordenador

Actualmente existe una serie de programas surgidos y desarrollados para ser utilizados exclusivamente por ordenador. En la década de los 90 se desarrollaron diversas utilidades de software en ámbitos clínicos, educativos y organizacionales. Por ejemplo, los conductores deben pasar un examen médico y otro psicológico para obtener y renovar sus licencias de conducir. En el contexto de la evaluación psicológica de los conductores se emplean algunos programas como el **DRIVERTEST** (Monterde y Bort, 1988)...., homologados por la Dirección General de Tráfico. Estos programas suelen incluir cuatro subtests: (1) Test de coordinación visomotora bidireccional a ritmo impuesto; (2) Velocidad de anticipación; (3) Tiempos de reacción múltiples; y (4) Atención concentrada y resistencia vigilante a la monotonía.

El test computerizado de relaciones espaciales complejas (RE-II) construido con el programa **HYPERCARD** para ser administrado y puntuado mediante Macintosh. Con un enfoque cognitivista, estudios realizados permiten considerarlo un instrumento útil en el campo de la inteligencia espacial (Fernández et al., 1992).

En evaluación clínica. Schundl y Bell (1993) han desarrollado un programa interactivo de ordenador que evalúa diferentes componentes de la imagen corporal. El Body Image Testing System **BITS**, presenta una vista frontal y de lado de un cuerpo humano y los sujetos pueden cambiar el tamaño de cada una de las 9 partes del cuerpo hasta que están satisfechos con la imagen creada. Este instrumento se ha utilizado para estudiar la relación entre distorsiones en la imagen del cuerpo y trastornos de la alimentación (Schundl y Bell, 1993).

En el contexto de la investigación clínica se ha desarrollado el **SYNWORK1**, programa que permite evaluar la ejecución de diversas tareas que implican memoria, procesamiento aritmético y motorización visual y auditiva, en un estimulado ambiente de trabajo. Este programa está siendo utilizado en experimentos de privación del sueño y desincronización

circadiana (Elsmore, 1994) y estudios llevados a cabo con el mismo han evidenciado que ciertos tipos de juegos de ordenador pueden mejorar el desarrollo de funciones cognitivas.

Gottschalk (1994) informa del desarrollo, validación y aplicación informatizada de métodos de medida de dificultades cognitivas e intelectuales a través del análisis del contenido de la conducta verbal. El procedimiento de análisis del contenido está basado en el método desarrollado por Gottschalk y Gleser para la medición de la magnitud de muchos otros estados y rasgos psicológicos además de disfunciones cognitivas. Gottschalk (1994) destaca la fiabilidad y objetividad del programa de inteligencia artificial recientemente desarrollado.

Así mismo, se han producido importantes desarrollos en las posibilidades de aplicación del ordenador en el ámbito educativo. Revisiones recientes de su utilización en ambientes educativos confirman la existencia de amplias posibilidades de utilización en diagnóstico educacional (Jäger, 1994). No obstante, Jäger (1994) subraya que muchos de estos programas se basan en el modelo-D, modelo que en la actualidad no ha resuelto todos sus problemas teóricos y metodológicos, requiriéndose una mayor elaboración teórica de los mismos.

En el ámbito organizacional, Schoenfeldt (1994) destaca que el número de aplicaciones por ordenador a una importante secuencia de actividades de la psicología organizacional (planificación, análisis de los puestos, selección...) se está incrementando en los últimos años (Schoenfeldt, 1994). El mayor impacto se ha planteado en la aplicación de tests y en la formación. En evaluación se utilizan para medir diversas variables como son las aptitudes mentales, las aptitudes psíquicas, aptitudes profesionales específicas, la personalidad y otras cualificaciones relacionadas con el trabajo (Schoenfeldt, 1994).

8.2.1.3. Aplicaciones combinando ordenador y vídeo

La combinación del ordenador con el vídeo ha hecho que nuevas habilidades más próximas a situaciones de la vida real puedan ser evaluadas, por ejemplo, la orientación espacial utilizando recorridos a través de esquemas y planos de ciudades, la comparación de objetos... (Zaal y Pieters, 1987). Recientemente, se ha utilizado el vídeo interactivo en orientación vocacional de personas sordas. El **SHIPS** permite medir habilidades interpersonales relacionándolas con la elección profesional en personas con esta dificultad (Loeding y Crittenden, 1994).

8.2.1.4. Tests adaptativos informatizados (TAI)

Un TAI (Test Adaptativo Informatizado) consiste en una estrategia informatizada de evaluación con la que se pretende estimar el nivel de habilidad (entendido en sentido amplio: como destreza, rasgo, aptitud, conocimiento...) de una persona de la manera más precisa y eficiente posible. Para realizar un TAI se necesita un banco de ítems calibrado, es decir, con propiedades psicométricas conocidas y obtenidas desde un modelo TRI y un algoritmo informático que controla la aplicación del test. A partir de ahí comienza la aplicación de un test a un sujeto por un ítem cualquiera (normalmente de dificultad media) y en función de las respuestas del sujeto el algoritmo va decidiendo el ítem siguiente a aplicar, hasta llegar al último, a partir del cual se estima el nivel de habilidad según la función de estimación asociada a ese ítem (Calero, 2007).

El desarrollo de la Teoría de Respuesta al Ítem (TRI) dio lugar a los tests adaptativos (Green, 1983). El diseño general de estas pruebas requiere necesariamente el uso del ordenador, ya que la aplicación a un sujeto es individualizada. La selección de los ítems que se aplican al sujeto depende de la exactitud de sus respuestas a los ítems previamente presentados. Por ello, es necesario disponer de un banco de ítems en el ordenador y de un programa informático mediante el cual, una vez seleccionado el ítem inicial de la aplicación, ésta se va adaptando a la ejecución del sujeto.

Son muchas las ventajas que se han señalado sobre este tipo de técnicas: entre otras, el aumento de la consistencia en la presentación de estímulos, la flexibilidad y precisión de su medida, el aumento de la amplitud del rango de la ejecución evaluada, la reducción del número de ítems que se aplican a un sujeto y el aumento de la motivación del evaluado al presentársele aquellos ítems que pueden ser ejecutados por él y eliminar de la aplicación los no accesibles a su nivel de ejecución (Green, 1983; Renom, 1993, 1994; Roid, 1986).

Según Wainer (2000) las ventajas de esta aproximación serían: 1) El aumento de la seguridad del test (pues este no se construye con una selección de ítems, sino con todos los ítems posibles); 2) La posibilidad de que el sujeto evaluado trabaje a su ritmo; 3) Que el sujeto se mantiene ocupado productivamente; 4) Que se resuelvan los problemas físicos de respuesta; 5) Que se puntúe inmediatamente; 6) Que los ítems se introduzcan en una secuencia lógica según la dificultad; 7) Que los ítems erróneos sean inmediatamente eliminados (los ítems menos discriminativos son los que menos utiliza el programa y van siendo eliminados); y 8) Que se puede incluir en un test una variedad de cuestiones diversas. A modo de ejemplo de TACs disponibles en España se pueden mencionar:

- Siguiendo este diseño se ha realizado una versión del test de Raven y de varias pruebas de rendimiento. En relación a los tests adaptativos computerizados para la medida del Factor “g”, Palacios, Colom, Espinosa y García (1994) presentaron en el IV Congreso de Evaluación Psicológica un análisis en relación a los mismos. El **TAC-g** es un instrumento de medida del factor “g” de inteligencia basado en dos fuentes principales: (1) a nivel teórico en los principios sobre “g” que estableció Spearman; y (2) a nivel instrumental en la Teoría de Respuesta al Ítem. Para operativizar los principios sobre “g” los autores construyen un banco de ítems con tareas verbales y numéricas de tres tipos: Tareas de educación de correlatos (ser opuesto de/ser la mitad de); Tareas de educación de relaciones basadas en propiedades físicas de los elementos (números) y semánticas (palabras); y Tareas de educación de relaciones y correlatos (analogías entre elementos). El modelo TRI que se usa para aplicar los distintos ítems se basa en la consideración de la dificultad-discriminación de dichos ítems y la estrategia de presentación es la de máxima información.
- **Morfo Bank** para la medición adaptativa de una habilidad lingüística. Adaptación castellana del Inventario Criterial del Lenguaje ICL (WIIG) a Test Adaptativo Computerizado (Renom y Puyuelo, 1994), y la evaluación adaptativa computerizada del razonamiento verbal consistente en la transformación del subtest de Razonamiento Verbal del TIR-4 a Test Adaptativo Computerizado (Renom, 1994).
- El **TRASI**, que es un test informatizado para la evaluación del razonamiento secuencial y la inducción (Rubio Franco y Santacreus, 2003).

- **La escala de ajuste emocional “EAE”**, incorporada como parte de un cuestionario de personalidad desarrollado a medida para fines de selección de personal (CPSP). La escala está compuesta por 28 ítems con seis opciones de respuesta en un gradiente de acuerdo-desacuerdo (Aguado et al., 2005). La revisión de la literatura acerca de la aplicación de los modelos de la teoría de la respuesta a los ítems (TRI) muestra que ésta se ha utilizado más en el trabajo con variables aptitudinales que para la medición de variables de personalidad, sin embargo diversos estudios han mostrado la eficacia de la TRI para la descripción psicométrica de dichas variables. Aún así, pocos trabajos han explorado las características de un Test Adaptativo Informatizado, basado en la TRI, para la medición de una variable de personalidad como es el ajuste emocional. Los resultados del estudio de Aguado et al. (2005) muestran la eficiencia del TAI para la evaluación del ajuste emocional, proporcionando una medición válida y precisa, utilizando menor número de elementos de medida en comparación con las escalas de ajuste emocional de instrumentos fuertemente implantados.

Recientemente, Wainer (2000) ha realizado un análisis del estado de la cuestión que dio lugar a los test adaptativos, centrando la discusión en el avance de los TAI y en particular en las expectativas que expresara Bert Green sobre las ventajas de esta tecnología. Analiza datos de la primera década donde los TAI fueron operativos, para comparar lo que ocurrió con aquello que se esperaba que ocurriera. En la misma dirección que Wainer (2000) podemos concluir que algunas de las metas de los TAI expresadas por Green están próximas a cumplirse, aunque la mayoría de ellas quedan para el futuro.

Como ha enfatizado otros investigadores (Wise y Kingsbury, 2000), los principios básicos de los tests adaptativos informatizados están relativamente bien establecidos. Sin embargo, la puesta en funcionamiento y el mantenimiento de un programa de tests adaptativos es bastante más complejo. Los profesionales de la medición habrán de enfrentarse a un conjunto de desafíos de tipo aplicado. En el trabajo se discuten desafíos de cuatro tipos: establecimiento y mantenimiento de los bancos de ítems, elección de los procedimientos de administración del test, protección de la seguridad del test, y respuesta a los asuntos relacionados con las personas que responden a los tests. El éxito de un programa de tests adaptativos dependerá mucho de cómo el elaborador de tests resuelva estos desafíos.

Desde los noventa, los test adaptativos computerizados han crecido en Estados Unidos de un modo exponencial y se utilizan hoy en exámenes de graduación, exámenes de admisión, pruebas de inglés para extranjeros y en la evaluación psicológica de las fuerzas armadas, y, su uso ha mostrado 2 ventajas adicionales: accesibilidad en puestos (muchos individuos pueden ser evaluados a la vez) y accesibilidad en tiempo. También se detectan algunas desventajas con respecto a la aplicación tradicional, a saber: mayor coste de los tests informatizados que los tradicionales (requieren una gran infraestructura), imposibilidad para el sujeto evaluado de volver atrás para revisar las respuestas dadas anteriormente y el problema motivacional derivado de la alta tasa de aciertos (superior al 50%) que el sujeto examinado tiene en esta situación test.

Para evitar este problema, en los últimos años ha surgido una versión nueva de tests autoadaptados informatizados (TADIs). La diferencia fundamental de estos con los anteriores es que el sujeto elige el nivel de dificultad en el que se quiere ubicar antes del comienzo y se le suministra información sobre el grado de acierto de cada respuesta. Como señala Calero

(2007), en España, a pesar de que existen varios grupos de investigación que trabajan en el diseño de este tipo de instrumentos, tal vez por la infraestructura requerida o quizás porque aquí no existen muchos contextos en los que se realicen evaluaciones masivas, hay pocos ejemplos de TAls.

8.2.1.5. Instrumento auxiliar en la aplicación de técnicas de potencial de aprendizaje

En la actualidad se utilizan como sistemas de evaluación-entrenamiento programas informáticos diseñados para suministrar al sujeto, de forma controlada, señales y reglas dirigidas a la solución de la tarea. Estos programas, que se basan en tareas clásicas tales como diseños con cubos, matrices progresivas, o memoria de palabras, muestran en pantalla el problema a realizar por el sujeto, dosifican las instrucciones y van suministrando reglas que facilitan la solución de la tarea y sirven al sujeto para generar una estrategia de solución de ese tipo de problemas. En concreto dos grupos de investigación utilizan este tipo de técnicas, el de Campione y Brown (1987) con niños, y el de Kliegl, Smith y Baltes (1990) con adultos.

En la década de los noventa, Gerber, Semmel y Semmel (1994) desarrollaron el **DynaMath**, un sistema de evaluación dinámica de multiplicación de múltiples dígitos mediante ordenador. Es un sistema a medida para estudiantes con dificultades. Dyna Math sistemáticamente organiza y da información sobre la realización de los estudiantes en problemas de multiplicación de diferente nivel de dificultad, mostrando gráficamente la zona de desarrollo proximal, destaca los tipos problemas a realizar en la siguiente fase de instrucción. El programa también registra automáticamente múltiples medidas de ejecución en una base de datos y provee un informe clínico con sugerencias para las intervenciones instruccionales.

Otro programa de ordenador de aplicación en ámbitos educativos es **EPA-TEST** desarrollado inicialmente por Calero y Araque (1990b) para evaluar el potencial de aprendizaje por ordenador, es decir, la modificabilidad de un sujeto que es sometido a un entrenamiento intelectual (**EPA-2**, Fernández-Ballesteros et al., 2000). Es un test de problemas matriciales tipo Raven, que utiliza para su presentación el ordenador. Está formado por 30 series de 3 problemas cada una, y sus posibilidades de uso son: como test estándar con tiempo limitado, como test estándar sin tiempo límite, como prueba de entrenamiento con tiempo limitado y como prueba de entrenamiento sin tiempo límite. El programa permite almacenar e imprimir los resultados obtenidos por distintos sujetos.

8.2.1.6. Desarrollo de Programas Informáticos Integrales

En la actualidad existen varios programas informáticos que integran diversas pruebas. Un ejemplo de estos, es el **PPD, Psychological Testing Program** (Zaal y Pieters, 1987) desarrollado a finales de los 80. Este programa incluye de 14 a 16 pruebas distintas que van desde tests perceptivos y verbales a cuestionarios de personalidad, valores interpersonales....., permitiendo diversos modos de presentación de la tarea, diversas tareas y diversos criterios normativos.

Otro paquete informático es el **Vienna System** (TEA, 1992) una herramienta de interés especial cuando la experiencia y la conducta humana son estudiadas con fines de diagnóstico o terapia. Algunas de las pruebas que incluye son: memoria visual, vigilancia, atención

continuada, detección de señales, rendimiento en series de trabajo, matrices progresivas, aptitud espacial, coordinación bimanual...

Otro programa diseñado es el **TEA System 4.0** (TEA, 2005), el cual integra diversos módulos en un solo sistema informático que permite el trabajo con los principales tests de evaluación de la personalidad (16PF, BFQ, CAQ, CTI, HSPQ, CPI, MMPI-2, MMPI-A, NEO-PI-R...) y de inteligencia (INFOWAIS III, INFOWISC-R, INFOWISC-IV...). Posibilita aplicar baterías de tests individuales, evaluar los resultados, definir los parámetros de cada test y ejecutar variadas funciones.

8.2.1.7. Aplicación de pruebas basadas en criterios (pruebas de competencia mínima)

Otro desarrollo en la utilización del ordenador para la administración de técnicas, se inició en la década de los 80 (Berts, 1986), consiste en el diseño y aplicación de pruebas basadas en criterios, denominadas en esta versión *pruebas de competencia mínima*. Estas pruebas permiten determinar la posición del individuo con respecto a un dominio de conducta previamente definida. En la actualidad existen pocas pruebas basadas en criterios.

8.2.2. APLICACIONES PARA LA CORRECCIÓN - PUNTUACIÓN DE PRUEBAS

En relación con la corrección e interpretación de pruebas, Fernández-Ballesteros y Calero (1992) subrayan tres desarrollos fundamentales.

1. En primer lugar, se ha facilitado la corrección mecánica de tests de plantillas, que actualmente se pueden realizar para gran cantidad de tests, mediante la utilización de hojas de respuesta especiales y lectoras ópticas. Un ejemplo de estos sistemas distribuido en nuestro país lo constituye el servicio de corrección mecanizada ofrecido por TEA.
2. Como un segundo desarrollo, se han incorporado algunos programas dirigidos a valorar y puntuar las respuestas de los sujetos a ciertas técnicas consideradas tradicionalmente de corrección conflictiva por la falta de estructuración de sus criterios valorativos, como son las técnicas proyectivas y algunos tests neurológicos. Existen programas para valorar las respuestas al test de Rorschach (con los sistemas de valoración de Klopfer y Exner), a tests proyectivos de dibujo, y también programas de valoración del test giestáltico visomotor de Bender y de la batería de Luria. La realización de estos programas puede suponer un gran avance en el grado de estructuración de los criterios de valoración que favorecerá un mayor acuerdo interjueces en relación con estas técnicas.
3. En tercer lugar, la contribución más importante del uso del ordenador para la corrección y puntuación de pruebas, está en la posibilidad de construir y aplicar un continuo normativo, o lo que es lo mismo, una especie de banco de normas, que permita la comparación de las puntuaciones obtenidas por un sujeto con cualquier criterio normativo o población de referencia previamente establecido.

Estas ventajas de elaboración de resultados no sólo han tenido lugar en relación a tests y autoinformes, sino también con técnicas tales como la observación y los registros psicofisiológicos.

Para facilitar la observación en situaciones naturales se ha construido el **Datacrhono** que es un aparato diseñado para el registro de datos de observación en situaciones naturales que permite el volcado posterior en un ordenador para su análisis e interpretación (Dpto. de Psicología Social de la Universidad Autónoma de Madrid y Dpto. de Bioingeniería de la Universidad Politécnica de Madrid) (ver Fernández-Ballesteros y Calero, 1992).

Otro sistema desarrollado y validado es el sistema de **evaluación ecoconductual (EBASS)** diseñado para realizar observaciones en el aula. Es un sistema observacional asistido por computadora que contiene 3 instrumentos (CISSAR, ESCAPE y MS-CISSAR) ampliamente utilizados en investigación en el campo de la educación especial. Los autores de la validación del instrumento (Greenwood et al. 1994) presentan las bases de estos sistemas en el contexto del creciente incremento de la necesidad de información sobre la conducta de los estudiantes y el ambiente de enseñanza. Así mismo, los autores confirman la eficacia del instrumento subrayando que el uso de sistemas computerizados puede permitir tomar mejores decisiones, ofreciendo mejores servicios y mejores resultados para los niños.

En el caso de **registros psicofisiológicos**, existen en nuestro país varios programas pilotos, por ejemplo, los desarrollados por Vila y Reyes para el análisis de la presión sanguínea (Reyes del Paso, 1991) en la misma dirección que la desarrollado en otros países (Demster, 2001).

8.2.3. APLICACIONES PARA LA INTERPRETACIÓN DE PRUEBAS

8.2.3.1. Tipos de informes computerizados

Existen programas de software que permiten realizar 4 tipos de informe a partir de los datos suministrados por el ordenador:

- *Informes descriptivos:* Se dan explicaciones sobre las significaciones de las escalas o de las elevaciones de las escalas.
- *Informes interpretativos:* Describen al sujeto de la evaluación, normalmente de forma narrativa, basándose en los resultados obtenidos por éste en los tests.
- *Informe integrado:* El informe contiene información de varias fuentes, varios instrumentos o varios autores.
- *Informe adaptado al usuario:* Este programa produce un informe desarrollado especialmente para el usuario o modificado a su medida sobre otro ya existente.

Fernández-Ballesteros y Calero (1992), en relación con la interpretación informatizada consideran tres tipos de programas:

- *Descriptivos.*
- *Basados en un modelo de predicción clínica.*
- *Basados en un modelo de predicción estadística.*

Los informes descriptivos han sido los más desarrollados. En éstos, se interpreta cada dato, variable o factor independientemente de los demás, pudiéndose obtener informes informatizados para el 16 PF, el MMPI o para las escalas de Wechsler. También se han elaborado programas que combinan síntomas psicopatológicos para establecer un diagnóstico

siguiendo el sistema de clasificación DSM-IV. Los programas informáticos desarrollados siguiendo un modelo de predicción sea clínico o estadístico, suponen hasta el momento el desarrollo más polémico. Existen escasos programas dirigidos a la realización del juicio clínico. La mayoría están centrados en el diagnóstico psiquiátrico o en la orientación escolar.

Cohen y Swerdlik (2001) diferencian los siguientes tipos de interpretación informatizada:

- *Informes de calificación (simples y/o extensos)*: Suministran puntuación por escala y perfiles. Los extensos informes de calificación proporcionan criterios de significación estadística de puntuaciones o de relaciones entre puntuaciones.
- *Informes de interpretación: Descriptivos, tamizados y/o consultivos*. Se acompañan de una narración que puede ser, en el primer caso, una descripción breve y concisa de cada una de las escalas. En el segundo incluye relaciones entre escalas, señalando aquellas que puedan ser interesantes para la interpretación. Los consultivos suministran un informe detallado en lenguaje adecuado para profesionales.
- *Informes integradores*: Tratan de combinar datos de diferentes fuentes, algunos incluso de diferentes modelos teóricos para su interpretación. Dentro de este grupo se pueden incluir los sistemas expertos.

8.2.3.2. Los sistemas expertos

Una aplicación del ordenador a la interpretación de datos evaluativos procede de los sistemas expertos. En el caso de los sistemas expertos, básicamente se trata de transferir los conocimientos y métodos de la inteligencia artificial a la resolución de problemas de diagnóstico y tratamiento.

- Un ejemplo de sistema experto, es el **CATCEC** (Computer-Supported Assessment and Treatment Consultation for Emotional Crises) de Hedlund, Vieweg y Cho (1987), programa de ayuda a la evaluación y tratamiento de las crisis emocionales.
- En España, Adarraga y Zaccagnini (1992) han construido el **DAI**, sistema que se ha mostrado fiable en el diagnóstico del autismo habiéndose elaborado a través de la formalización, mediante el lenguaje computacional, de las estrategias para el diagnóstico de un experto en esta patología.
- En una línea, hasta cierto punto similar, Westmeyer y Hagebock (1992) han elaborado dos programas de ordenador (**PSYMEDIA** y **DIASYS**) a través de los cuales formalizan el proceso evaluativo y, por tanto, permiten auxiliar, mediante un modelo normativo, el trabajo del evaluador a través de tal proceso. El PSYMEDIA es un sistema de inferencias que ayuda al evaluador a elaborar un diagnóstico con datos de pruebas basadas en la teoría clásica de los tests. El DIASYS 1 es un sistema aplicable a la evaluación psicoeducativa.
- En el campo educativo, se ha desarrollado un sistema experto para identificar procedimientos de tratamiento para estudiantes con dificultades severas. El **SMH: PAL** emplea la tecnología de los sistemas expertos para ayudar al profesor a identificar intervenciones que pueden ser utilizadas con estudiantes que exhiben determinados problemas de conducta (Hofmeister et al., 1994).
- Otro sistema de este tipo han sido el **IPFSE**: Intelligence Psychological Disorders Symptoms Evaluator, que sigue el formato de diagnóstico del DSM-IV (Cheng y Liou, 2003).

Los sistemas expertos, ayudan al evaluador en la tarea de toma de decisiones diagnósticas en base a los datos obtenidos en el proceso de evaluación. Son programas informáticos que se acercan a la forma de razonar de los expertos humanos. En contraste con los modelos estadísticos (aplican un modelo algorítmico al proceso diagnóstico ayudado por el ordenador) que utilizan datos cuantitativos y algoritmos para tomar decisiones, los sistemas expertos usan la inferencia y el razonamiento. Por ello su uso hace posible tratar con problemas sobre los que no se puede conseguir una información completa, o que estén insuficientemente estructurados, o donde la base de conocimientos sea susceptible de modificación, aspectos todos ellos presentes, de forma necesaria, en la evaluación con seres humanos. La aplicación de estos sistemas al diagnóstico psicológico supone una ayuda de gran valor, puesto que a esta flexibilidad de razonamiento humana se une la mayor capacidad del ordenador en cuanto a memoria, confiabilidad y capacidad de procesamiento. En los últimos años se está trabajando en la construcción de sistemas de este tipo aplicados a áreas muy específicas.

Las características funcionales de un sistema experto o KBS (Knowledge-Based System) son las siguientes (Calero y Fernández-Ballesteros, 2004):

- Es capaz de ejecutar tareas difíciles que normalmente requieren para su realización un humano experto.
- Puede ser completado recurriendo a métodos heurísticos y dominios de conocimiento mejor que con algoritmos eficientes (estos algoritmos no son necesarios en la ejecución de un experto).
- La tarea tiene un cierto grado de definición objetiva o intersubjetiva en términos de categorías básicas implicadas, metas, criterios y conocimiento aplicado. Los métodos de trabajo puramente subjetivos son excluidos.
- Se espera que el sistema trabaje bajo condiciones de cierta incertidumbre.
- Puede ser capaz de explicar y justificar su comportamiento.
- Es indispensable el uso apropiado de interfaces.

Las características estructurales son (Calero y Fernández-Ballesteros, 2004):

- Una base de conocimientos que contiene una representación del dominio: a) de hechos (descripciones de situaciones establecidas o hipótesis sobre las que el sistema va a razonar; y b) de reglas (definidas por expertos y que corresponden a operaciones significativas en el dominio de conocimientos).
- Un motor de inferencias: parte informática del sistema que provee de las fuentes de procesamiento necesarias para aplicar el conocimiento.
- Una memoria que almacena datos y conclusiones del problema particular.
- Un interface que incluye facilidad de justificación y explicación.

A partir de su transformación a lenguaje informático, las funciones asignadas a los sistemas expertos son las siguientes (Calero y Fernández-Ballesteros, 2004):

- Facilitar la expresión de reglas, que harán lo más próximas posible a las formuladas por los expertos.

- Explotar la unión de reglas: a) combinar o agrupar las reglas para deducir nuevos conocimientos, decisiones, predicciones o juicios...; y b) poder rendir cuentas acerca de cómo se llega a conclusiones.
- Dar la posibilidad de añadir, modificar o suprimir reglas.

Una de las peculiaridades de los sistemas expertos es la separación del lenguaje informatizado del lenguaje experto. En psicología son fundamentales para traducir conocimientos poco estructurados, interesantes por las capacidades de evolución del sistema y útiles por su potencial pedagógico. Además, permiten la integración de diversos modelos la utilización en contextos diversos y la incertidumbre de los resultados.

La calidad de los sistemas expertos es un tema de estudio relevante en la actualidad, preocupando a distintos estudiosos las técnicas y métodos que se emplean para dar validez y fiabilidad a los mismos. La evaluación de los sistemas expertos es un tema de gran relevancia, pero como señalan Sharma y Conrath (1993) en ocasiones las técnicas de verificación y validación más comúnmente señaladas en la literatura suelen ser insuficientes. Estos autores realizan una revisión de los métodos y subrayan la necesidad de un método común para el desarrollo de los sistemas expertos. Sharma y Conrath (1993) analizan tres métodos, el Turing-Test (más ampliamente utilizado, se considera que la meta del sistema experto es minimizar la conducta humana, muestra dificultades de generalización en ambiente natural), los métodos cuantitativos (más restrictivos y simples, limitados en sus perspectivas, no proveen una impresión conjunta de calidad) y los métodos híbridos. Desde su perspectiva, los enfoques que parecen tener mayor amplitud para evaluar los sistemas expertos son los métodos híbridos que emplean criterios cuantitativos y cualitativos mediante técnicas de evaluación subjetiva.

En 2009 se han cumplido 58 años de la publicación del fascinante ensayo: ¿Puede pensar una máquina? de Alan Turing. Los actuales desarrollos en los campos de las tecnologías de la información, la inteligencia artificial y la ciencia cognitiva nos permiten, lanzar otra pregunta: ¿puede una máquina enseñarnos a pensar? Hoy podemos sugerir que los sistemas expertos pueden desempeñar esta función, ya que “después de aprender él mismo, puede enseñar, no solo hechos y datos sino también procedimientos” que pueden ser facilitadores de nuevos procesos de pensamiento.

8.2.3.3. Programas informáticos en el mercado español

Existen diversos productos informáticos distribuidos en el mercado español. Algunos representativos de éstos son:

- **Info-WAIS** (Calvo Mauri, 1993): Programa que facilita la corrección y elaboración de informes de la escala WAIS. A partir de la introducción de las puntuaciones directas se obtienen diversos índices y se elabora automáticamente el perfil del sujeto y un breve informe.
- **Info-WAIS III** (Seisdedos, Fluiters y Carrasco, 2004). Programa informático en Windows para ayudar a la corrección e interpretación de la escala WAIS-III.
- **RR**. Sistema experto para la interpretación del Test de Rorschach (Drag & MCF, 1995; García Polín, 1996): Es un sistema concebido como ayuda para la interpretación del test de

Rorschach. Basado en un potente sistema experto, aprende con los sucesivos casos y se maneja fácilmente.

- **RECORD.** La técnica de la Rejilla (Feixas y Cornejo, 1992): Es un instrumento informatizado que permite evaluar las dimensiones y estructura psicológica del constructo "significado personal", aspectos de la forma en que una persona da sentido a su experiencia y organiza su mundo. En el contexto de la orientación vocacional se emplean programas de ayuda para este asesoramiento. Diez, Gironella y Feixas (1993) presentan un caso de fracaso escolar en el que se realiza un asesoramiento vocacional mediante la técnica de la rejilla que permite analizar las implicaciones y dilemas cognitivos que supone la elección de profesión para el adolescente. La rejilla se analiza mediante el programa informático RECORD (Feixas y Cornejo, 1992) que facilita la aplicación de la técnica y su cómodo análisis mediante complejas técnicas multivariadas que sugieren una imagen gráfica del significado personal que facilita su análisis. Estos autores ilustran la forma de proceder desde la perspectiva constructivista en el ámbito psicoeducativo y de intervención psicológica
- **PPI** (Sebastian y Machlab, 1994): Es un programa informático para evaluación de la personalidad a partir del test 16 PF. Dispone de una base de datos interactiva y elabora un perfil de la personalidad y un completo informe sobre la misma y sus posibles desajustes, incluyendo coeficientes de correlación con otros casos.
- **Info-WISC-R** (Seisdedos, 1993): Programa de ayuda a la corrección e interpretación de la escala WISC-R. Se introducen las puntuaciones directas obtenidas por el sujeto y el ordenador ofrece las puntuaciones transformadas de acuerdo con el manual, perfil del niño, y un informe para uso del psicólogo con los comentarios pertinentes.
- **SOCGRAMA** (Almari y Gil, 1993): Consiste en un programa informático para el tratamiento de los sociogramas que, hasta ahora, resultaba muy laborioso a la hora de tratar y tabular los datos.
- **PSIQUE CP** (PSICWARE, 1993): Es un sistema informático para gestión de un gabinete de Psicología Clínica con varios terapeutas. Permite realizar la gestión de los pacientes, conservar historias clínicas, y la aplicación de diversas pruebas de evaluación de uso frecuente en el campo clínico (EPQ, STAI, STAIC...), así como la obtención de informes sobre los resultados del sujeto.
- **METRIX Engine** (Renom, 1996): Sistema de análisis y evaluación de tests. Facilita la gestión e interpretación de resultados en bases de datos y aporta las bases metodológicas y técnicas para la elaboración de tests de todo tipo.
- **CPTII.** Test de atención de Conners. Evaluación computerizada de la atención
- **VISIONLAB:** Sistema de gafas con obturador conectadas a un ordenador que permite la presentación de estímulos visuales en tres dimensiones. El ordenador puede funcionar como taquistoscopio, tambor de memoria analizador de percepción visual...
- **Test de golpeteo digital computerizado:** Test de golpeo utilizando el teclado del ordenador para medir la coordinación.
- **DATALAB2000:** Dieciséis canales de evaluación de variables fisiológicas: ECG, EMG, EEG, GSR, presión sanguínea, conducción nerviosa, tensión muscular, ritmo cardíaco, respiración, fuerza, estimulación muscular...
- **Sistema Poligráfico Computerizado:** Cuatro canales para la evaluación fisiológica: respiración, respuesta galvánica de la piel, presión sanguínea/ritmo cardíaco.

- **BIOPAC:** Sistema modular de biofeedback y control fisiológico de doble canal.
- **BEI.** Batería Escolar Informatizada: batería de pruebas para edades de 6 a 18 años, que evalúa aptitudes intelectuales, personalidad, técnicas de estudio, intereses profesionales... (TEA, 2008)

8.2.4. IMPACTO Y POLÉMICA DE LA EVALUACIÓN INFORMATIZADA

La gran preocupación mundial por la incursión del ordenador en el quehacer evaluativo, se pone de manifiesto con la reciente realización de guías y estándares para regular su uso (Bartram y Coiné, 2003), que se describieron en el capítulo 6. La polémica sobre el uso de ordenador en Evaluación psicológica se plantea, por un lado, entre los psicólogos que argumentan a favor o en contra de su utilización y, por otro lado, entre los que asumen la facilitación del ordenador en la administración de pruebas, pero rechazan los informes interpretativos que en ocasiones ofrecen.

8.2.4.1. La administración informatizada

En relación al empleo de los ordenadores en la evaluación psicológica, los defensores de la misma destacan el hecho de ahorrar tiempo al profesional, disminuir costes, e incrementar la fiabilidad por la mayor estandarización en la toma de datos. En esta dirección, Urbina (2007) subraya que una manera eficaz de evitar errores en la aplicación de los tests y el riesgo de que el resultado se vea influido indebidamente por variables relacionadas con el sexo, la raza, la edad, la apariencia o el estilo personal del examinador es la aplicación del test mediante el ordenador. Sin embargo, otro grupo de psicólogos manifiestan recelos frente a la administración de entrevistas y tests por ordenador. Entre quienes ven con inquietud la generalización de su uso en la administración se puede diferenciar dos cuestionamientos muy disímiles: a) La crítica basada en la necesidad de contar con una mayor verificación empírica de ambas versiones del mismo test; y b) la crítica desde una posición humanista sustentada en la despersonalización de la tarea.

Ejemplo de esta segunda controversia centrada en el rol es la posición de Tallent (1987) quien enfatiza que la administración computerizada desvirtua la complejidad del proceso al transmitirle al sujeto una imagen de la evaluación en lo que lo central es el instrumento y no el grado de compromiso y la responsabilidad del psicólogo en su relación personal con el entrevistado. Sampson (1990) responde a esto afirmando que no se produciría una verdadera despersonalización ya que el contacto humano nunca se pierde porque siempre debe de haber una persona que supervise y provea de asistencia a fin de asegurar la validez de los resultados.

Una fuente de temor de un grupo de profesionales está relacionada con la creencia de que la función del psicólogo puede ser suplantada por el ordenador. Golden (1987) haciendo referencia a la evaluación neuropsicológica informatizada, plantea un recelo frente a su utilización al observar que no está claro por qué los psicólogos se muestran tan ansiosos por utilizar estos programas, y que si los tests pudieran ser interpretados por ordenadores, obviamente habría una necesidad mucho menor de psicólogos. Sin embargo, desde un punto de vista diferente, Fowler y Butcher (1986) consideran que el ordenador no puede suplantar al clínico, ni está pensado para hacerlo, de la misma forma que el ordenador de abordo no puede

suplantar al piloto. En ambos casos, puede liberarles de cargas tediosas y añadir nuevas dimensiones de rapidez, exactitud y economía en las tareas que tendrán que realizar.

Algunos autores (Sampson, 1990) alertan contra caer tanto en la tecnofilia como la tecnofobia: "Si la computación va a enriquecer la evaluación dependerá más de las actitudes y capacidades del profesional que de los desarrollos de los programas. El descubrimiento del fuego por los humanos hubiera sido una nota insignificante en la historia si no hubieran encontrado una aplicación efectiva a la nueva tecnología" (Sampson, 1990). Una posición semejante es la de Weiner (1991) quien refiriéndose a su uso en la investigación indica que puede ser "héroe" o "villano" ya que puede, según el caso, favorecer investigaciones cruciales o producir una enorme cantidad de datos muy precisos respecto a cuestiones irrelevantes.

Así mismo, Wetzler y Marlowe (1994) reflexionando sobre el papel del ordenador en la evaluación clínica consideran que puede ser de ayuda en relación a componentes técnicos de la psicología clínica, pero que no puede hacer el trabajo de evaluación psicológica e intervención. Desde su punto de vista los ordenadores son herramientas y como cualquier herramienta sólo es adecuada si lo es su usuario.

Los trabajos que han comparado el uso del ordenador para *administrar autoinformes* con la tradicional situación de aplicación, confirman que no se aprecian diferencias en cuanto a la cantidad de información recogida, ni en variables como la deseabilidad social en los cuestionarios de personalidad y actitudes (Booth-Kewley, Edwards y Rosenfeld, 1992). No obstante, algunos investigadores recomiendan adaptar los formatos para la aplicación por ordenador (Watson, Thomas y Anderson, 1992).

En relación a la calidad algunos autores afirman que la administración por ordenador parece aumentar la percepción de confidencialidad y conducir a una menor cantidad de respuestas sesgadas (Esformes y Geva, 1987; Skinner y Allen, 1983). Así mismo, Honaker y Fowler (1990) al referirse a la evaluación informatizada, aunque admiten que aún requiere ajustes para hacerla menos estática, subrayan como una de sus ventajas que, por su carácter impersonal, puede favorecer un mayor sinceramiento en sujetos muy sensibles a quienes les resulta más fácil admitir sus dificultades a una máquina que planteárselas al entrevistador.

Recientes estudios españoles (Olea, Revuelta, Ximénez, y Abad, 2000) sobre los efectos psicométricos y psicológicos de las respuestas en tests fijos y adaptativos informatizados, aplicaron dos versiones informatizadas de un test de vocabulario inglés para hispanohablantes (uno fijo y otro adaptativo) a una muestra de estudiantes españoles de primer curso de Psicología. Se estudiaron los efectos del tipo de test (fijo versus adaptativo) y de la condición de revisión (permitida versus no permitida) sobre diversas variables psicológicas. Se analizaron los efectos de la revisión en ambos tests (diferencias antes-después) en una serie de variables psicológicas y psicométricas. Después de la revisión, dos fueron los resultados más destacables: a) un incremento significativo del número de aciertos y de la habilidad media estimada; y b) un descenso significativo del nivel de ansiedad-estado de los evaluandos. No se obtuvieron diferencias significativas en precisión. Tampoco resultó significativo el efecto de la interacción entre el tipo de test y el momento (antes versus después de la revisión). En el trabajo que estos investigadores llevaron a cabo se discuten estos resultados y se comentan las implicaciones que los resultados pueden tener para permitir la revisión en la aplicación real de tests adaptativos informatizados.

Durante las dos últimas décadas muchos estudios han demostrado la capacidad de las entrevistas administradas por ordenador para recoger información clínicamente significativa de

pacientes adultos en el contexto de servicios de salud mental, y más recientemente esta investigación se ha extendido al ámbito clínico con niños y adolescentes. Algunos de los estudios que han investigado los efectos de proveer a los clínicos con un informe computerizado de entrevistas previas a la evaluación clínica realizada por ellos con niños han concluido que la disponibilidad de estos informes influyó en el tipo de problemas identificados por los clínicos y los servicios que ellos recomendaron para tratar los problemas de los niños (Sawyer, Sarris y Baghurst, 1992). A partir de estas conclusiones se sugiere la necesidad de realizar más investigación que estudie los beneficios de la utilización del ordenador en los servicios de salud mental en niños y adolescentes. En este sentido, la revisión realizada por Kaplan y Saccuzzo (2006) sobre la administración computerizada de entrevistas les lleva a concluir que las personas son más francas y sinceras al responder a una computadora que a un terapeuta experimentado.

Recientemente, Proyer y Häusler (2007) han llevado a cabo un estudio teórico de los tests objetivos de personalidad computerizados y los muestran como una posible forma de evaluar el comportamiento en situaciones concretas. Desde su punto de vista, una de las principales características de estos tests es que no presentan validez aparente y son, por consiguiente, menos susceptibles de falseamiento u otras distorsiones de respuesta que habitualmente se encuentran en la evaluación subjetiva de personalidad mediante cuestionarios. En este trabajo examinan los desarrollos recientes en el área de tests de personalidad objetivos, analizan su aplicación con muestras clínicas y no clínicas y los tests objetivos de personalidad se discuten como alternativas en el campo de evaluación psicológica, como instrumentos válidos y económicos para la evaluación del comportamiento.

Las investigaciones que han analizado los factores que pueden limitar el uso de los ordenadores en evaluación destacan como factores relevantes: 1) La falta de entrenamiento adecuado y falta de experiencia del evaluador (Farrell, 1989); y 2) Una actitud negativa frente a los ordenadores (Farrell, 1991) o una actitud ansiosa frente al uso de ordenadores por parte de los evaluados (Fariña, 1993).

En relación a la administración computerizada, con respecto a las posibilidades de almacenamiento de datos, las ventajas del ordenador son múltiples, pues se acumulan datos de diferentes sujetos, se construyen baremos propios, se tienen en cuenta patrones de respuesta que resulten de interés... En este sentido la revisión realizada por Kaplan y Saccuzzo (2006) sobre la administración de pruebas por computadora les conduce a concluir que este tipo de administración generalmente consume menos tiempo, es más económica, es mejor aceptada por los examinados ya sean adultos o niños y con frecuencia es más precisa.

Si las ventajas son bastantes, las limitaciones también pueden ser importantes, y de éstas quizás la más cuestionada sea la relativa a la equivalencia de formatos. Cuando utilizamos una versión informatizada de un test clásico, se considera que es equivalente y que evalúa lo mismo que la versión de papel y lápiz. Sin embargo, cuestiones como el límite de tiempo o el forzar la "no respuesta" influyen en los resultados obtenidos. Otros autores destacan como limitación que el uso de estos soportes puede promover, en los usuarios, actitudes pasivas hacia la evaluación psicológica y un aumento de la accesibilidad a los tests por parte de clientes potenciales que podría provocar un entrenamiento inadecuado (Snyder, 2000). Una limitación importante de la administración computerizada es que no permite captar las sutilezas de la comunicación humana, el lenguaje corporal. Aunque algunos programas se están volviendo sofisticados al grado de poder percibir la emoción y ajustar las respuestas como corresponde (Picard y Klein,

2002), aún tienen que igualar la capacidad del psicólogo para detectar claves emocionales sutiles (Kaplan y Saccuzzo, 2006).

En general muchas investigaciones aportan datos sobre una actitud positiva y la percepción de una mayor confiabilidad por parte de los sujetos evaluados, aunque algunos estudios han mostrado, como efecto colateral no deseado el aumento de las puntuaciones y la aparición de perfiles más patológicos (en concreto en la versión informatizada del MMPI).

8.2.4.2. La interpretación informatizada

En este punto la controversia se da, por un lado, entre quienes opinan que la computadora solo debe generar perfiles e interpretaciones descriptivas a modo de hipótesis, dejando la construcción del informe final en manos del psicólogo y, por otro lado, entre quienes apoyan los informes interpretativos narrativos. Estos informes narrativos incluyen recomendaciones terapéuticas y para la devolución, siendo presentados generalmente con una copia para el psicólogo y otra para el cliente.

Respecto a los informes interpretativos realizados por el ordenador se han sugerido numerosos problemas. Entre otras cuestiones, la validez de la narración generada por el ordenador como juicio clínico sobre un sujeto en particular no ha sido establecida, y además, este tipo de descripciones suelen ser tan generales que podrían ser aplicadas a infinidad de sujetos (Skinner y Pakula, 1985).

De hecho, en 1986 la APA publicó una guía "*Guidelines for Computer-based Tests and Interpretations*", subrayando que "los informes interpretativos generados por un ordenador sólo podrán ser usados junto con el juicio de un profesional" (APA, 1986). En estas directrices si bien se acepta la situación de evaluación informatizada como similar a la aplicación de técnicas en formato de papel y lápiz, siempre que se comprueben las propiedades psicométricas de esas versiones, sin embargo, se cuestionan tanto las clasificaciones como los informes interpretativos realizados mediante ordenador.

Entre quienes cuestionan el uso de los informes narrativos se dan dos tipos de argumentos, aquellos que sostienen que la interpretación computerizada aún no ha logrado su madurez científica, y aquellos quienes piensan que además atenta contra la centralidad del rol del psicólogo en la tarea de integración de los datos de la evaluación.

Matarazzo (1986) enfatiza que la aceptación acrítica de un informe computerizado implica un cierto enajenamiento por parte del psicólogo. Este renuncia así a su propio conocimiento y juicio clínico, a la posibilidad de integrar el material jerarquizando los aspectos que considera más significativos y a la capacidad de comunicarlos en su estilo personal, para hacerse en cambio eco de un pensamiento y de una voz, ajenos. Dejaría entonces de ser un clínico experto para convertirse en un simple técnico de la evaluación. Así mismo Groth-Marnath en 1990, y Frank de Verthelyi (1996) refrendan esta postura subrayando la importancia de la capacidad de discriminación e integración del clínico: "el psicólogo no debe aceptar nunca ciegamente un informe narrativo, sin asegurarse previamente de que las afirmaciones incluidas estén basadas en datos de investigación recientes, concuerden con sus propias observaciones y tengan coherencia dentro del contexto de la historia y la situación particular de cada paciente".

Ambos autores comparten la creencia de que aunque el informe incluya un aviso que indique las limitaciones de una interpretación probabilística siempre existe el peligro de su uso inadecuado por personas que aceptaran de forma poco crítica el contenido del informe

narrativo. En 1992, esta postura crítica y cautelosa frente al informe computerizado también es expresada por Anastasi en un artículo sobre el uso de los tests, en el que destaca que el error estaría en creer que el informe computerizado es necesariamente certero, cuando en realidad es un elemento auxiliar (Anastasi, 1992).

Además, y de acuerdo con Frank de Verthelyi (1996) la cualidad de un autoinforme depende de un número de factores, incluidos la validez del test en sí mismo, la corrección de las reglas de decisión que sustentan el sistema interpretativo y la cualidad y el número de afirmaciones interpretativas con que cuenta el sistema. Así, algunos autores que han analizado informes del MCMI destacan que son muy útiles para la descripción de las características más estables de personalidad, pero pueden presentar déficits serios en cuanto a los síndromes clínicos y al diagnóstico.

8.2.4.3. Limitaciones actuales de la evaluación informatizada

Actualmente existen gran cantidad de tests informatizados de todas las técnicas que se han descrito en el capítulo anterior: entrevistas, cuestionarios, tests de inteligencia, de aptitudes, de intereses, técnicas proyectivas, observación, registros psicofisiológicos.... Sin embargo algunos estudios han puesto de relieve que su uso permanece constante desde el inicio de los noventa. Farrell (1989) evidenció que los psicólogos de la APA en los años ochenta utilizaban el ordenador para la administración de técnicas en un 26%, para la puntuación en un 41% y para la interpretación en un 29%; y una década más tarde, McMinn et al. (1999) hallaron porcentajes similares.

¿Qué razones pueden explicar la ausencia de incremento en la incorporación del ordenador al quehacer evaluativo?

- 1) *Falta de entrenamiento y formación:* Algunos psicólogos, han señalado como causa la falta de formación de los psicólogos en el uso de tests computerizados. Farrell (1989) señaló que la falta de un adecuado entrenamiento, aprendizaje y experiencia con la tecnología podría ser la causa que explicara que los psicólogos en ejercicio no utilizaran las nuevas tecnologías. Trabajos recientes (Olson, 2001) han evidenciado que sólo un 25% de los programas de doctorado y masters de Estados Unidos incluyen cursos de formación sobre tests informatizados.
- 2) *Coste económico:* Otra razón que puede ser importante es el coste económico que supone mantener los equipos y programas actualizados.
- 3) *Dudas sobre la equivalencia de la administración tradicional e informatizada:* La desconfianza y la falta de control que sobre la toma de decisión percibe el psicólogo evaluador parecen ser también razones para no utilizar el ordenador en evaluación. Sin embargo, aunque diversos trabajos señalan la equivalencia entre tests informatizados y tests de papel y lápiz (Merten y Ruch, 1999, Bressani y Downs, 2001, ver Kaplan y Saccuzzo, 2006), otros autores señalan que existen factores específicos derivados de la administración que pueden dar lugar a la no equivalencia. Es más, algunos muestran variabilidad intraindividual del efecto (Gantt y Green, 1986), en el sentido antes apuntado de que el formato informatizado puede dar lugar a puntuaciones más altas.

- 4) *Resultados incongruentes de la interpretación*: También se constata que un ordenador convierte variables continuas en discretas, asignando un valor de interpretación a cada una de las categorías, por lo que, a veces, se produce una incongruencia entre interpretaciones (Zinder, 2000). Además, es bien conocido que la evaluación informatizada deja fuera otros elementos no verbales que en una evaluación tradicional se recogen como señales: gestos faciales, tono de voz, etc., y anulan la interacción evaluador-evaluado que puede tener cierta influencia en la ejecución (Butcher, Perry y Atlis, 2000). No obstante, el uso del ordenador en administración y puntuación, aún con reservas, es, hoy, aceptado.
- 5) *Problemas de validez*: Se han hecho tres tipos de estudios sobre validez de la evaluación informatizada (Snyder, 2000):
- Los centrados en el grado de satisfacción del cliente: algunos han señalado una aceptación positiva por parte de los evaluados, pero otros indican que las interpretaciones informatizadas pueden ser menos individualizadas y excesivamente generales (Knauss, 2001) y que en la mayoría de ellas se abusa de los descriptores no diferenciales, hasta tal punto que se advierte que el uso del ordenador aumenta el “efecto Barnum” en los informes (descripciones en las que se mezclan estereotipos, vaguedades y evasividad, incluyendo afirmaciones universalmente válidas).
 - Los que se refieren al uso de modelo de expertos: en general se ve como una buena estrategia inicial para garantizar la validez de un programa el que un experto realice un modelo previo, aunque no se considera suficiente (Zinder, 2000).
 - Los focalizados en el uso de criterios externos: estos son los realmente demostrativos de validez, pero hasta la fecha no han arrojado resultados contundentes. Las muestras de validación del formato informatizado suelen ser pequeñas, y las medidas externas de criterio suelen ser inadecuadas (Knauss, 2001). Existen escasas investigaciones sobre la consistencia de los informes, e incluso algunas de éstas detectan discrepancias entre informes, lo que pone de relieve los problemas de validez.

Como han señalado Calero y Fernández-Ballesteros (2004), a pesar de las ventajas del uso del ordenador (su precisión, confiabilidad y eficiencia, la economía, inmediatez, eliminación de errores humanos, posibilidad de recopilación sistemática de datos y de aplicación a poblaciones especiales), pero debido a una serie de puntos débiles (el coste, la falta de validez demostrada y la falta de control sobre una decisión de la que el evaluador tiene la última palabra), hacen que el uso de esta tecnología no haya despegado a nivel aplicado tanto como se esperaba.

Además, Calero y Fernández-Ballesteros (2004), subrayan algunos problemas éticos de la evaluación informatizada indicando que hay que tener en cuenta: 1) que una evaluación basada en la tecnología computacional es siempre de orientación normativa; 2) que los datos de los clientes en un ordenador no están protegidos adecuadamente; y 3) que, evidentemente, su uso tiene un efecto diferencial en función de ciertas variables personales tales como el efecto de la práctica o la ansiedad que genere en el sujeto. Por todo ello se requiere un doble análisis de su implantación: el económico (costo/beneficio) y el deontológico (considerando la adecuación de las normas profesionales generadas por esta nueva situación).

Respecto a las futuras posibilidades de la evaluación por ordenador, éstas se encuentran relacionadas con diversos factores como son: el desarrollo de la tecnología, la progresiva

miniaturización de los ordenadores, su abaratamiento, la simplificación en la interacción con el ordenador así como la formación de los profesionales..., y la superación de algunas de las limitaciones señaladas previamente.

Desde la década de los 90 existen ordenadores que no necesitan teclado, y que se utilizan escribiendo directamente sobre la pantalla con un lápiz y papel, y este tipo de procedimientos puede resultar de una gran ayuda en la observación y la autoobservación. Otra línea en desarrollo actual son los ordenadores capaces de reconocer la voz humana y gestos humanos, que serán accesibles en el mercado dentro de unos pocos años, siendo este desarrollo de gran ayuda en la evaluación de personas con limitaciones físicas, visuales o auditivas. Así mismo el mundo de las simulaciones a través de ordenadores y videos interactivos o a través de la más compleja realidad virtual puede ayudar en la evaluación de conductas en situaciones a la vez realistas y controladas.

Para concluir se puede señalar que teniendo en cuenta las ventajas que proporciona la evaluación informatizada, es esperable que superadas las limitaciones que actualmente tiene, en el futuro se incorpore en el quehacer profesional del psicólogo con mayor frecuencia.

8.3. LA EVALUACIÓN DE PROGRAMAS, INTERVENCIONES O TRATAMIENTOS

La evaluación de programas es un campo que se encuentra en continua expansión. La primera etapa de la historia de la evaluación de programas (1965-1975) estuvo caracterizada por el interés en el logro de alta objetividad (Scriven y Campell). El segundo período (1975-1982) supone una reacción contra el uso del modelo científico, y se llama la atención sobre la importancia del contexto sociopolítico en la implementación de programas. En esta etapa se aceptan diseños orientados al análisis del proceso, a la búsqueda de resultados más allá de la contrastación de las hipótesis planteadas, y se emplean métodos cualitativos. El tercer período (1980-actualidad) es una época de síntesis que está interesada en hallar el punto de unión entre las divergencias señaladas, buscando métodos de análisis o de diseños que puedan ser aplicados sin perder la objetividad (Cook, 1988).

En la evaluación de programas en contextos educativos, se observa la misma evolución, ya que se han sucedido cronológicamente 4 modelos evaluativos (Pérez Juste, 2006): 1) las propuestas experimentales que evalúan la eficacia de los programas, su resultado, es decir, si ha conseguido sus objetivos, debiendo utilizar para ellos medidas o pruebas con garantías psicométricas de fiabilidad y validez; 2) los modelos orientados a los procesos que surgen tras observar que los diseños experimentales no ayudan a formular mejoras de los programas, siendo su valor fundamentalmente formativo, analizando especialmente los factores que afectan a la eficacia del programa; 3) los modelos complejos que tienen en consideración otras cuestiones además del análisis del resultado y del proceso, tales como circunstancias de partida, los medios y recursos disponibles, la disponibilidad de personal, las expectativas, las exigencias de los destinatarios, la adecuación del programa al proyecto educativo del centro..., ya que desde este modelo se observa que un mismo programa puede ser eficaz en unos contextos u otros en función de diversas variables; y 4) los modelos cualitativos, que tienen como principal interés no tanto la valoración de los resultados como la descripción exhaustiva de los programas para comprender y desentrañar la compleja realidad educativa, primero utilizan la observación, después seleccionan cuestiones relevantes del programa en su contexto

de referencia, y por último intentan explicar los principios subyacentes a la organización del programa tratando de establecer las relaciones causa-efecto.

La cantidad de modelos existentes y su heterogeneidad indica que, por ahora, se está muy lejos de una conceptualización única acerca del concepto de evaluación de programas. Los distintos modelos muestran una gran disparidad metodológica, algunos defienden una alta objetividad en el análisis con el uso de diseños y medidas de orden experimental, otros utilizan una evaluación de tipo naturalista y de bajo o nulo control experimental, mientras que otro grupo de modelos defienden una síntesis de ambas perspectivas.

8.3.1. BASES CONCEPTUALES

8.3.1.1. DEFINICIÓN Y OBJETO DE ESTUDIO

Como ya se ha señalado en el primer capítulo, la Evaluación Psicológica puede tener por objetivo producir un cambio en el comportamiento, y por lo tanto exigir una intervención (tratamiento). En estos casos la evaluación (variable/s dependiente/s) se realiza antes y después del tratamiento (variable independiente), y sus resultados permiten valorar el tratamiento aplicado.

Este particular tema, el del ámbito de la evaluación de intervenciones o valoración de programas supone una nueva aplicación, una nueva extensión de la Evaluación Psicológica. Aunque evaluación y valoración son sinónimos en castellano, algunos autores consideran que la *evaluación* es un término descriptivo y al término *valoración* se le otorga una carga semántica de enjuiciamiento y valoración del cambio. En este trabajo, y siguiendo la misma dirección que diversos autores (Anguera, 1994a; Arruabarrena, 2009; Arruabarrena y De Paul Ochotorena, 2002; Campo y López, 2006; Cava y Musitu, 2000; Fernández-Ballesteros, 1995b; Forns, 1990, 1993; Forns y Gómez, 1995; Garaigordobil, 1992ab, 1994c, 1995ace, 1996ab, 1999ab, 2000, 2001, 2002, 2004ab, 2005abc, 2006, 2008ab, 2009, Garaigordobil y Berrueco, 2007; Garaigordobil y Echebarria, 1995; Garaigordobil, Maganto y Etxeberria, 1996; Garaigordobil, Maganto, Pérez, y Sansinenea, 2009; Martínez, Pérez, Navarro y Redondo, 2007; Garaigordobil y Pérez, 2004; Martorell, 1988b; Lameiras, Carrera, Núñez, y Rodríguez Castro, 2006; Sánchez y Anguera, 1993), los emplearemos como sinónimos.

La evaluación de intervenciones no sólo implica el juicio sobre concretos tratamientos realizados a concretos individuos en demanda de tratamiento y cambio de conducta. La implicación del psicólogo no se reduce a la valoración de psicoterapias individuales sino que, además, colabora en el enjuiciamiento de otras intervenciones que, con objetivos de cambio, se realizan en el amplio contexto social. Por ejemplo, programas educativos, sanitarios, etc. Así, el psicólogo evaluador no sólo participa en la evaluación y/o valoración de concretas intervenciones individuales sino que actúa también, dentro de un equipo multiprofesional cuando el objetivo es valorar intervenciones sociales que tienen por finalidad producir un cambio en la conducta.

Con el término *evaluación de programas*, inmerso en un programa de Evaluación Psicológica, estamos haciendo referencia a la sistemática investigación del mérito, valor o éxito de concretos tratamientos, intervenciones o programas llevados a cabo con el fin de producir cambios comportamentales. El objetivo de tales intervenciones sería resolver problemáticas sociales. En este contexto, el objeto de evaluación es un tratamiento, intervención o programa

implantado en la realidad social a través de los cambios comportamentales producidos (Fernández-Ballesteros, 1992g). Aunque Fernández-Ballesteros en sus primeros trabajos (1985, 1992g) matiza las diferencias entre el concepto evaluación (individuo) y valoración (programa) en sus escritos posteriores (Fernández-Ballesteros, 1995b, Fernández-Ballesteros, 2004f), y en la misma dirección que otros autores, habla de evaluación cuando se miden los efectos de intervenciones realizadas en contextos educativos y/o terapéuticos.

Pese a que el término valoración de programas fue acuñado para hacer referencia a la evaluación de programas concretos de acción social promovidos por los poderes públicos, la valoración de intervenciones se desenvuelve también en ámbitos individuales del psicólogo aplicado que evalúa la efectividad de su propio trabajo. Aunque tienen sus matices diferenciales, es obvia la similitud que la valoración de programas guarda con la valoración de un tratamiento en el caso individual en el que se produce una demanda de cambio. En ambos casos, resulta necesario definir los objetivos, clarificar el conjunto de actividades y acciones en que consistirá la intervención, y seleccionar los instrumentos de medida adecuados para evaluar el cambio y realizar posteriormente una valoración del mismo. Una revisión del concepto de uso de “valoración de programas” ha sido realizada por Esteve et al. (1993).

En líneas generales evaluar un programa consiste en recoger información sistemática sobre las operaciones implicadas en ese conjunto de acciones encaminadas a un objetivo (Hernández y Rubio, 1992). Así, definimos la evaluación de programas como la sistemática investigación a través de métodos científicos de los efectos, resultados y objetivos de un programa con el fin de tomar decisiones sobre él (Fernández-Ballesteros, 1995b). Usualmente dos conceptos o juicios valorativos aparecen asociados a la evaluación de un programa, la eficacia del programa y la efectividad. Se dice que un programa es eficaz cuando ha conseguido las metas u objetivos que se proponía, mientras que se dice que ha sido efectivo cuando con independencia de los objetivos establecidos el programa ha tenido efectos positivos.

Para Stufflebeam, uno de los autores pioneros más influyentes en este campo, es especialmente importante que la evaluación contribuya al perfeccionamiento de los programas y la comprensión de los fenómenos estudiados, dado que la complejidad de unos y otros exige una mejora constante de los modos de enfrentarnos con los problemas. En sus propias palabras “la evaluación es el proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y mérito de las metas, la planificación, la realización y el impacto de una actuación determinada con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados” (Stufflebeam y Sinkfield, 1987).

Con el fin de que la evaluación pueda contribuir a la consecución de los tres propósitos señalados, Stufflebeam propone un modelo que implica evaluar: 1) La adecuación de las decisiones relacionadas con la selección de las metas y objetivos, decisiones que van a contextualizar el planteamiento y la realización de cualquier programa de intervención; 2) La adecuación de las decisiones relacionadas con la selección de los métodos y procedimientos, esto es, de las acciones que se han de poner en marcha para conseguir los objetivos buscados; 3) La adecuación de la aplicación y desarrollo del programa, esto es, el proceso concreto de actuación, proceso que puede diferir de lo proyectado en la fase anterior; y 4) Los efectos derivados de la aplicación del programa, esto es, el producto obtenido tras el desarrollo del mismo, producto que puede deberse o no a los métodos seleccionados (Alonso Tapia, 2004b).

Las situaciones aplicadas en las que tienen lugar los programas de intervención son muy variadas. Anguera (1994a) en la conferencia impartida en el IV Congreso de Evaluación Psicológica realiza una revisión de los artículos publicados en los últimos 5 años en las revistas *Evaluation Review*, *A Journal of Applied Social Research*, y *Evaluation & The Health Professions*, perfilando la siguiente panorámica: Programas comunitarios, para el ámbito de la salud, en el ámbito escolar, recursos humanos, recursos universitarios, programas epidemiológicos, programas de atención a personas sin hogar, programas preventivos de adicciones, de embarazos en adolescentes, de asistencia a las familias, programas penitenciarios, programas para mujeres maltratadas.... entre otros. Así, la expresión evaluación de intervenciones en situaciones aplicadas es enormemente abarcadora. No obstante, debe diferenciarse la evaluación de programas de tratamiento de la evaluación de estrategias o políticas sociales (Urnau, 1993).

8.3.1.2. CRITERIOS PARA LA EVALUACIÓN DE PROGRAMAS

Ante todo un programa debe cumplir 4 criterios: la factibilidad o viabilidad, la utilidad, la eticidad y la precisión o exactitud (Pérez Juste, 2006). En 1981, el *Joint Committée on Standards for Educational Evaluation*, publicó el "*Standards for Evaluations of Educational Programs, Projects and Materials*", en el que se proponen 30 normas agrupadas en 4 categorías o criterios: 1) Utilidad: Este criterio incluye 8 normas destinadas a asegurar que la evaluación proporcione la información que necesite cada uno de los receptores legítimos; 2) Viabilidad: Contiene 3 normas concebidas para promover una evaluación realista y prudente; 3) Honradez: El criterio integra 8 normas destinadas a asegurar que la evaluación sea legal, ética, equitativa...; y 4) Precisión: Incorpora 11 normas técnicas destinadas a asegurar la calidad de la información sobre la realidad objeto de estudio, de tal modo que sea posible determinar su valor o mérito. Algunos años más tarde, en 1994, se realizó una revisión de estos criterios que fue publicada en "*The Program Evaluation Standards*".

En el área de la evaluación de programas se ha introducido el concepto de evaluación de la evaluabilidad, un concepto clave en la Teoría de la Evaluación de Programas. Se trata de analizar en qué medida un programa o servicio es evaluable como paso previo al inicio de la evaluación del mismo. Por evaluabilidad se entiende la facilidad con la que un programa puede ser bien evaluado (Fernández-Ballesteros, Hernández et al., 1990). La medida en la cual un programa puede ser evaluado, depende en buena parte de hasta qué punto ha sido bien planificado así como hasta qué punto pueden existir barreras que dificulten la evaluación (Fernández-Ballesteros, 1995d). La evaluación de la evaluabilidad ejerce una doble función, por un lado, la clarificación de si la evaluación es factible, y por otro lado, sirve de control previo de las cuestiones relevantes del programa o servicio. El análisis de la evaluabilidad permite predecir en qué medida esa evaluación permitirá la extracción de información y la posterior valoración del programa. Esta evaluación de la evaluabilidad se establece como un proceso, dentro del de planificación, diseño y evaluación de programas, que permite enfrentarse al hecho de lo costoso de la evaluación de un programa (Hernández y Rubio, 1992).

Con la finalidad de valorar la evaluabilidad del programa se elaboró un "*Listado de Cuestiones relevantes en Evaluación de Programas LCREP*" que permite analizar el grado de definición de 14 dimensiones básicas de una evaluación de programas o servicios: Evaluación de necesidades; Establecimiento de objetivos; Definición del programa; Calidad de los criterios de selección; Nivel de implantación; Calidad del diseño; Calidad de las operaciones; Calidad del

procedimiento de recogida de información; Posibilidad de recogida de información sobre el contexto; Aceptabilidad; Obstáculos; Implicación del evaluador; Especificación de la finalidad de los resultados de la evaluación; y Costos de la evaluación (Fernández-Ballesteros et al, 1990, 1995). En opinión de Forns (1990) este listado recoge muchas de las cuestiones problemáticas acerca de la evaluación de programas... pero cabe plantearse si el LCREP es utilizable para todo tipo de programas o si se ajusta con mayor precisión a algún tipo de ellos; por ejemplo, difícilmente una evaluación judicial puede ser evaluada mediante el LCREP.

Sanders (1998) identifica 4 estándares en la evaluación de programas: 1) la utilidad, desde la que se deben contemplar los siguientes aspectos: la identificación de los interesados, la credibilidad del evaluador, el foco y selección de la información, la identificación de valores, la claridad del informe, la oportunidad y difusión del informe, y el impacto de la evaluación; 2) la viabilidad que tiene en cuenta los procedimientos prácticos, la viabilidad política, así como la eficacia de los costes; 3) la propiedad que tiene en cuenta algunos aspectos tales como la orientación hacia el servicio, los derechos de las personas, las relaciones humanas, la revelación de los resultados o conflictos de intereses; y 4) la precisión, desde la que se tienen en cuenta cuestiones como la documentación del programa, el análisis del contexto, los propósitos y los procedimientos descritos, las fuentes de información adecuadas, la información válida, fiable y sistemática, así como el análisis de la información tanto cuantitativa como cualitativa. En síntesis, en evaluación de programas, la evaluación debe cumplir 4 criterios: 1) debe ser útil, proporcionando la información práctica necesaria; 2) debe ser realista, es decir, factible de llevar a cabo; 3) debe ser ética; y 4) debe ser precisa (Garaigordobil, 1998).

8.3.1.3. MODELOS EN LA EVALUACIÓN DE PROGRAMAS: EVALUACIÓN SUMATIVA Y FORMATIVA

La valoración de un tratamiento implantado a un sujeto o a un grupo se juzga en función de haber conseguido los cambios de conducta predichos y demandados. Sin embargo, existen distintas matizaciones en la valoración de intervenciones, pudiéndose diferenciar dos tipos de evaluación: formativa y sumativa. La evaluación formativa es aquella que se realiza sobre el proceso de implementación de una determinada intervención, tratamiento o programa, mientras que la evaluación sumativa se refiere a los resultados derivados de la aplicación de una intervención. Scriven (1967) establece las diferencias entre evaluación sumativa y formativa según una serie de aspectos que se exponen en el Cuadro 42.

Cuadro 42. Características de la Evaluación Formativa y Sumativa

	EVALUACIÓN FORMATIVA	EVALUACIÓN SUMATIVA
Roles	Mejorar un programa de estudios en curso de elaboración (valoración de acciones formativas).	Determinar el valor de este programa cuando su implementación ha sido acabada (balance final).
Personas que realizan la evaluación	El evaluador debe ser preferentemente el propio implementador del programa. Precisa de un evaluador-implementador altamente conocedor del programa en el aspecto cognitivo y emocional.	El evaluador debe ser una persona externa e independiente del proceso que se va a evaluar.
Momento evaluativo	Evaluación continua: Se realiza a lo largo y en distintos y sucesivos momentos de la implementación del programa con la	Evaluación final: Se considera el producto al final, teniendo por finalidad la sanción social o determinación del

	finalidad de ejercer la función de corrección del proceso.	grado de precisión o éxito global alcanzado.
Tipo de información que se analiza	Tendencialmente realizan análisis moleculares con tendencia a ignorar los efectos locales de los programas.	Puede ser más molar, proporciona información descriptiva y tiende a enfatizar los efectos locales del programa.

Analizando la evaluación de programas en el contexto educativo, Forns (1993) destaca que la acción formativa ejercida en un proceso de evaluación puede adquirir un aspecto estructural o un aspecto constructivista. Con esta dicotomía se pretende diferenciar, en el dominio de la evaluación de aprendizajes, entre dos tipos de evaluación formativa. Una se centraba en la estructura intrínseca del propio conocimiento transmisible y otra en los niveles o sucesión de estadios personales de organización de un conocimiento, denominando estos enfoques respectivamente evaluación formativa conductual o estructural y evaluación formativa constructivista. Esta autora (Forns, 1990) clasifica los programas en función de cuatro grandes objetivos:

1. *La valoración de programas centrada en el análisis de la consecución de objetivos:* La clave del análisis es el logro de los objetivos previamente definidos. La principal crítica a este modelo es que la valoración no puede surtir efecto de perfeccionamiento sobre lo evaluado ya que se realiza en una fase final, ante el producto acabado.
2. *La valoración de programas centrada en la aportación de información como medio de sustentar la Toma de Decisiones:* Una valoración orientada a la toma de decisiones supone el análisis del proceso o, en su caso, del resultado de la aplicación de un programa con la finalidad de determinar los índices críticos sobre los cuales fundamentar la selección entre varias posibilidades de acción. La valoración debe proveer juicios de valor basados en indicadores claves que permitan decidir si un programa debe interrumpirse, modificarse en algunos aspectos, o continuar. Dependiendo de quién sea el responsable de la toma de decisión se puede diferenciar la valoración orientada a la toma de decisiones por parte del beneficiario, y la valoración orientada a la toma de decisiones por parte de las instancias administrativas.
3. *Valoración de programas centrada en el enjuiciamiento de sus efectos generales:* Bajo este epígrafe pueden ubicarse la evaluación libre de metas de Scriven, la evaluación adversaria de Owens y la evaluación judicial de Wolf. Desde esta perspectiva se defiende un tipo de valoración en la que predomina el descubrimiento de las relaciones causales entre los distintos acontecimientos, o el establecimiento de nexos de relación entre los hechos finales y sus antecesores, para lo cual es imprescindible un análisis exhaustivo de la información. En la evaluación libre de metas, el evaluador desconoce las metas del programa, a fin de que no vea influido su juicio por la referencia a las finalidades u objetivos perseguidos. En la evaluación adversaria y judicial el eje de la valoración se centra en la discusión de los aspectos clave del programa, desde puntos de vista diversos que son atacados y defendidos por evaluadores distintos. Se parte de la creencia de que los evaluadores difícilmente pueden ser observadores no-sesgados, por lo que, la participación en el programa de diversos equipos de evaluadores entrenados a discutir los puntos de vista de sus adversarios, puede ayudar a esclarecer puntos cruciales.

4. *La valoración de programas centrada en el uso social de los resultados de la evaluación:*
Este modelo de evaluación se centra en el uso social de los resultados de la evaluación.

Otra clasificación de los modelos fue planteada por Anguera (1994) para la que existen dos grandes modelos en evaluación, que son el lineal y no lineal, estrechamente relacionados con las denominadas evaluación sumativa y formativa, respectivamente, así como con la evaluación como resultado y como proceso. El primero de estos macromodelos parte de la secuencialidad de las tres fases de planificación, implementación y evaluación de un programa, lo cual equivale a decir que la evaluación (medida de los efectos de un programa) depende rigidamente de su estructuración y agenda inicial, sin apenas posibilidad de modulación en función de circunstancias no previsibles que derivan del carácter complejo que presenta cualquier intervención. La característica definitoria de una evaluación de resultados se reduce estrictamente a estimar la magnitud de los efectos de un tratamiento sobre el cual se había establecido previamente una hipótesis a modo de conjetura sobre dichos previsibles efectos. El macromodelo no lineal, por el contrario, se basa en la continua interacción y modulación recíproca entre planificación, implementación y evaluación, lo cual equivale a la intersección continuada entre los tres componentes de relevancia, progreso y eficiencia. La evaluación entendida como proceso pretende entender por qué se ha producido un determinado resultado; la finalidad no es demostrativa, sino explicativa, y en este enclave será más cómodo situar la factibilidad de la evaluación.

En síntesis, en la evaluación sumativa (evaluación del cambio, del resultado) se estima magnitud de los efectos del programa en función de las hipótesis, tiene una finalidad demostrativa, el evaluador debe ser externo e independiente, y en este modelo de evaluación la planificación, evaluación, e implementación están estructuradas desde el principio. En la evaluación formativa (evaluación del proceso) se intenta entender porqué se ha producido determinado resultado, la finalidad es explicativa, resulta conveniente que el evaluador sea el que implementa el programa, y en ésta se produce una continua interacción y modulación recíproca entre planificación, implementación y evaluación. En función de los objetivos que se pretendan conseguir con la evaluación se opta por un modelo u otro, aunque en ocasiones y en algunos aspectos ambos modelos pueden utilizarse de forma complementaria.

Finalmente conviene comentar que el uso de evaluadores internos o externos es una de las polémicas omnipresentes en el ámbito de la evaluación de programas. Los detractores del valorador interno siempre han aducido en contra su ausencia de objetividad. Sin embargo, algunos estudios que analizan los factores que influyen en la credibilidad de los programas (Esteve et al., 1995) concluyen que, aparte de la objetividad, a la hora de juzgar la credibilidad de las valoraciones, hay otros valores en alza, como por ejemplo, su cercanía al programa. En este estudio los resultados evidenciaron la relevancia de factores como las credenciales del equipo de valoración (universidad, ministerio, equipo del barrio...) y el modelo de práctica valorativa denominado modelo de síntesis (modelo que defiende que la selección de objetivos y metodología debe tener en cuenta la naturaleza del problema valorativo concreto). Así, una fuente experta y digna de confianza incrementa la credibilidad de la evaluación.

A este respecto Fernández-Ballesteros (1995b) examina las ventajas e inconvenientes de la evaluación desde dentro (realizada por los mismos profesionales que ejecutan el programa) y de la evaluación desde fuera (evaluador ajeno al programa y su implantación) destacando las ventajas que se exponen en el Cuadro 43.

Cuadro 43. Ventajas e inconvenientes de la evaluación interna o externa

	VENTAJAS	INCONVENIENTES
Evaluación desde dentro	• Optima la mejora del programa. • Minimiza la reactividad de los sujetos. • Menos costosa. • Mayor influencia sobre el programa.	• Minimiza la objetividad. • Minimiza la credibilidad social de la evaluación. • Minimiza la utilización de estándares y tecnología dura.
Evaluación desde fuera	• Maximiza la objetividad. • Maximiza la utilización de tecnología dura (diseño, instrumentos...). • Maximiza la credibilidad social de la evaluación. • Maximiza la utilización de estándares.	• Minimiza las posibilidades de mejorar el programa. • Maximiza la reactividad de los sujetos. • Más costosa. • Menor influencia sobre el programa.

Tal vez las ventajas de un evaluador interno o externo pueden estar relacionadas con el tipo de programa objeto de evaluación, por lo que en cada caso convendrá examinar la bibliografía referente al impacto de esta variable en el tipo de programas que vaya a ser objeto de evaluación.

8.3.1.4. TAREAS Y FASES DEL PROCESO EN LA EVALUACIÓN DE PROGRAMAS

La evaluación de programas se lleva a cabo mediante un proceso de toma de decisiones a través del cual se planifica, se examina, se recogen datos y se informa sobre el valor o mérito del objeto de evaluación. El proceso de evaluación/valoración ha sido estudiado desde muchas perspectivas, y diversos autores han puesto de relieve las etapas básicas del proceso, enfatizando como es lógico unas sobre otras.

Martorell (1988b) destaca entre las tareas básicas a realizar en evaluación de programas las siguientes:

1. Determinación del propósito del programa: Para la determinación de los propósitos es necesario tener conocimiento de: (1) con qué tipo de organización o ayuda se cuenta; (2) a qué tipo de organización va dirigida la intervención; (3) sobre qué personas va a recaer el programa; y (4) metas que se pretenden.
2. Clarificación de a quién va dirigido.
3. Articulación del programa.
4. Determinación de lo que va a ser objeto de valoración: Quién es el staff, quién o quienes son los patrocinadores y solicitantes, que actividades hay que realizar, a qué tipo de personas va a ser aplicado, con qué recursos se cuenta, de qué supuestos se parte.
5. Especificación de metas y efectos.
6. Especificación de las variables contextuales.
7. Especificación de variables intervinientes.
8. Fijar el "cómo" y "cuando" debe de llevarse a cabo la evaluación.

Otros investigadores (Anguera, 1994) subrayan los siguientes requisitos básicos de planteamiento a dilucidar en la proposición de un programa:

1. Elección del modelo de evaluación: Elección de un modelo de evaluación sumativa (lineal), y/o un modelo de evaluación formativa (no lineal).
2. Características del programa en cuanto a sus objetivos, actividades previstas y plan de acción general.
3. Recursos necesarios y recursos disponibles para la implementación del programa, tales como, recursos personales, instrumentales, temporales, económicos, fungibles, etc.
4. Factibilidad o evaluación de la evaluabilidad, es decir, análisis de en qué medida el programa puede ser bien evaluado.
5. Principios éticos que compendian la tarea de los evaluadores tales como autonomía, buena fe, contribución al bienestar del otro, justicia y fidelidad.

Respecto a las cuestiones básicas a tener en cuenta a la hora de planificar un programa de intervención Alonso Tapia (2004), propone preguntarse 13 cuestiones así como las implicaciones de las mismas:

1. ¿Qué sujetos van a participar en el programa?: Implica definir las características que hacen relevante la participación en el programa.
2. ¿Cuántos sujetos van a participar en el programa?: Implica determinar los sujetos a atender. La cifra tiene implicaciones tanto económicas como organizativas.
3. ¿Quiénes van a aplicar el programa?: Implica determinar los recursos humanos con que se cuenta y las necesidades de personal.
4. ¿Qué necesidades de formación tienen las personas que han de aplicar el programa?: Implica buscar recursos personales, materiales y económicos para la formación de las personas que realizan la intervención.
5. ¿Qué condiciones debe tener el plan de formación de las personas que han de aplicar el programa?: Las personas que deben recibir formación precisan que ésta tenga ciertas características para ser efectiva, características que es preciso definir y asegurar de antemano.
6. ¿En qué va a consistir el programa?: Implica determinar las características del entrenamiento –qué se va a enseñar, cómo se va a enseñar, durante cuánto tiempo, con qué materiales... Se requiere que la intervención esté adecuadamente fundamentada desde un punto de vista teórico.
7. ¿Por qué ese programa y no otro?: Implica considerar las ventajas e inconvenientes de las alternativas de actuación cuando las hay.
8. ¿Qué personas se requieren para gestionar el programa?: Depende de la envergadura del programa.
9. ¿Dónde, cuándo y durante cuánto tiempo se va a aplicar el programa?: Es fundamental que la temporalización y la duración del programa sean las adecuadas para dar tiempo a que se produzcan los cambios deseados.
10. ¿Cómo garantizar que el programa se aplica según el plan previsto?: Puede que quienes han de aplicar el programa no lo hagan del modo adecuado si no se les incentiva adecuadamente.
11. ¿Cómo valorar de modo preciso tanto el desarrollo del programa como su impacto?: El diseño del proceso a seguir para valorar el programa debe constituir un elemento esencial del proceso de planificación.
12. ¿Qué financiación requiere?: Es algo que deriva de los restantes puntos

13. ¿Cómo conseguir la financiación?: Es fundamental para que el programa sea viable.

Para Fernández-Ballesteros (Fernández-Ballesteros, 1992g, 1995c; Fernández-Ballesteros y Fernández del Valle, 2003) seis son las fases fundamentales del proceso de valoración-evaluación de intervenciones:

1. *Planeamiento de la evaluación*: Objetivos y especificación del programa y de su implementación.
2. *Selección de operaciones a observar*: Variables dependientes que van a ser medidas u observadas a través de determinados instrumentos.
3. *Establecimiento del diseño de investigación*: Selección del diseño adecuado y análisis de variables que podrían contaminar los hallazgos.
4. *Recogida de información*: Logística de la recogida de datos, aplicación de instrumentos y mecanismos de control de las variables contaminadoras que puedan estar implicadas en el proceso de aplicación de tests y otros dispositivos de medida.
5. *Análisis de datos*: Análisis descriptivos o inferenciales en función del objetivo de valoración, el diseño y la naturaleza de los datos cuantitativos o cualitativos recogidos.
6. *Informe valorativo*: Producto de la valoración que va a permitir la toma de decisiones sobre la intervención.

En cuanto a las etapas y cuestiones a abordar en el proceso de valoración de un programa Alonso Tapia (2004) subraya 7 etapas:

- Etapas:
- Etapas 1. Focalizar el estudio de valoración: ¿Quién demanda el estudio de valoración y para qué? ¿Cuáles son las preguntas cruciales a las que se debe buscar respuesta? ¿Qué valorar? ¿Quién puede verse implicado o afectado por el estudio de valoración? ¿Qué factores de la situación pueden afectar a la valoración? ¿Es posible llevar a cabo el estudio de valoración?
 - Etapas 2. Decidir la dirección, organización y coordinación de la valoración: ¿Quién debe dirigir y realizar el trabajo de valoración? ¿Qué costo económico puede tener? ¿Cómo organizar y temporalizar el trabajo de valoración? ¿Qué problemas cabe esperar que surjan?
 - Etapas 3. Diseñar la valoración: ¿Qué modelos alternativos pueden servir para diseñar el estudio de valoración? ¿Qué procedimientos utilizar para decidir el diseño de la valoración? ¿Qué elementos incluir en el diseño del estudio de valoración?
 - Etapas 4. Seleccionar instrumentos y recoger la información: ¿Cómo recoger el máximo de información con el mínimo costo? ¿Qué clases de información se deberían recoger y quién debe recogerla? ¿Qué procedimientos utilizar para recoger la información que se necesita? ¿Es necesario desarrollar instrumentos o basta con seleccionar entre los existentes? ¿Cómo determinar la validez y fiabilidad de los instrumentos?
 - Etapas 5. Analizar e interpretar la información: ¿Cómo organizar los datos? ¿Cómo analizarlos? ¿Qué criterios se deben utilizar para interpretar los resultados del análisis?
 - Etapas 6. Elaboración de informes: ¿Es necesario que alguien reciba un informe del estudio de valoración? ¿Qué contenidos deberían incluirse en el informe? ¿Cómo comunicar los datos del informe? ¿Cuál debe ser la estructura y el estilo del informe? ¿Cómo ayudar a quien lo reciba a interpretar y usar su contenido?

Etapas 7. Evaluación de la valoración: ¿Es preciso evaluar la calidad del trabajo de valoración?
¿Quién debe evaluar el trabajo de valoración? ¿Qué criterios utilizar para evaluar la valoración?

En similar dirección, Álvarez González et al. (2006) han propuesto las siguientes fases de un programa en contextos educativos: 1) El análisis del contexto ambiental donde se va a desarrollar el programa; 2) Identificación de necesidades; 3) Diseño del programa (fundamentación teórica, formulación de objetivos, contenidos que se deben desarrollar, selección de actividades para desarrollar, recursos para el desarrollo de las actividades, temporalización de la realización de las actividades, destinatarios, criterios de evaluación de cada objetivo y costes); 4) Ejecución del programa; y 5) Evaluación del programa. En opinión de estos autores para poner en marcha un programa se debe partir de una serie de requisitos mínimos: 1) el programa ha de ser asumido por el centro, que debe contar con personas que lo dinamicen; 2) existen unos mínimos que todo centro debe garantizar (coordinación, implicación, preparación, recursos, ayuda técnica) si se quiere implantar la intervención por programas; 3) disposición de recursos suficientes dentro y fuera del aula; y 4) formación de los agentes del programa.

Álvarez González et al. (2006) conceptualizan varias fases en la evaluación de programas: 1) Elaboración del plan de evaluación (objetivo de la evaluación, quién recibirá los resultados, quiénes se encargarán de realizar la evaluación, qué se tiene que evaluar, cuando se tiene que evaluar, cuáles son los criterios de la evaluación); 2) Instrumentación y recogida de la información (qué conviene tener en cuenta para la selección de instrumentos, qué tipo de instrumentos y recursos podemos utilizar); 3) Análisis y valoración de la información (qué aspectos se tendrán en cuenta en el análisis, qué fases incluye el análisis y valoración de los resultados); y 4) Presentación de los resultados (respondiendo al propósito de la evaluación, e indicando todos los aspectos que debe tener en cuenta un informe de evaluación).

Para Pérez Juste (2006), en la evaluación de programas se debe llevar a cabo: 1) una evaluación inicial, en la que se evalúan las necesidades existentes y en base a la cual se identifican los objetivos del programa, que permita valorar la adecuación del programa a los destinatarios, al contexto donde se aplica el programa, su adecuación a las circunstancias (viabilidad)...; 2) una evaluación formativa del programa, a través de la que se analizan las actividades, las secuencias, la temporalización, los recursos, la flexibilidad... y 3) una evaluación final a través de la que se miden los resultados, los efectos del programa, su eficacia, su impacto.

En un intento de ordenar de forma sencilla el proceso y las tareas que llevamos a cabo cuando diseñamos y evaluamos un programa, en este proyecto docente sistematizamos este proceso enunciando 4 grandes fases: 1) diseño del programa, 2) diseño de la evaluación, 3) implementación del programa y de la evaluación del mismo, y 4) elaboración del informe de la evaluación del programa. En cada fase se debe reflexionar y dar respuesta a diversas cuestiones que se describen a continuación.

FASE 1. DISEÑO DEL PROGRAMA DE INTERVENCIÓN

En el diseño de un programa de intervención han de determinarse las siguientes cuestiones:

1. *Justificación y antecedentes del programa:* Razones por las que se va a llevar a cabo, interés del programa, ventajas-utilidades del programa que va a diseñarse y antecedentes del programa.
2. *Fundamentación teórico-empírica del programa:* Características y necesidades en la etapa evolutiva para la que se está diseñando el programa; definición de conceptos teóricos relevantes al programa; explicaciones teóricas propuestas por las distintas perspectivas epistemológicas de la psicología sobre los conceptos teóricos; factores que influyen en las variables de intervención; investigaciones previas similares o complementarias, así como procedimientos y resultados obtenidos en ellas.
3. *Objetivos del programa:* Identificar objetivos generales que se pretenden con el programa así como objetivos específicos, concretos incluidos en estos objetivos generales.
4. *Características del programa:* Determinar a quién va dirigido, donde se va a realizar, quién lo administrará, durante cuanto tiempo...
5. *Configuración del programa:* Identificar los módulos del programa y el número de actividades que contiene cada módulo o sección, las técnicas que subyacen a las actividades y al programa en su conjunto...
6. *Actividades o acciones del programa:* Las actividades constituyen la esencia de un programa por lo que se requiere describir detalladamente las actividades seleccionadas para cubrir los objetivos, las actividades que configuran el programa deben describirse de forma homogénea, por ejemplo, en formato de ficha técnica (objetivos de la actividad, descripción de la actividad e instrucciones para los participantes, preguntas para la discusión o debate posterior a la actividad, materiales necesarios para ejecutarla...).
7. *Procedimiento para la implementación del programa:* Redactar la consigna o encuadre de introducción al programa, elaborar un cuadro o esquema que determine la temporalización de la secuencia de las acciones a realizar para aplicar el programa, tiempo que durará la implementación, número de sesiones de intervención que se llevarán a cabo, tiempo de duración de cada sesión, actividades que se realizarán en cada sesión...
8. *Factores de influencia en el programa:* Factores o variables que pueden afectar positiva o negativamente a la implementación del programa, por ejemplo, la formación de los adultos que implementarán el programa, el espacio físico donde se realice...

FASE 2. DISEÑO DE LA EVALUACIÓN DEL PROGRAMA

El diseño de la evaluación de un programa requiere dar respuesta a las siguientes cuestiones:

1. *Finalidad de la evaluación y factibilidad:* En primer lugar, el evaluador debe clarificar la finalidad de la evaluación del programa, si la evaluación que se propone es factible de llevarse a cabo y qué dificultades podría encontrarse durante la evaluación.

2. *Identificar las variables dependientes y los instrumentos de evaluación con los que se medirán estas variables:* En este punto el evaluador debe preguntarse, en primer lugar, qué variables dependientes se van a medir teniendo en cuenta la finalidad de la evaluación, cuales deben ser las variables objeto de análisis para evaluar el programa que permitan mejor evaluar los cambios, y si existen otros efectos del programa que pueden ser evaluables. En segundo lugar, debe identificar con qué instrumentos de evaluación medirá las variables dependientes seleccionadas como objeto de evaluación. Deberá valorar si existen instrumentos estándar con garantías psicométricas de fiabilidad y validez o se requiere la creación de instrumentos “ad hoc” para medir estas variables. Así mismo se preguntará por las fuentes que ofrecen mayores garantías para la recogida de datos, es decir, quién aportará la información de mayor calidad para la evaluación de los efectos del programa. Se deberá determinar el procedimiento a seguir para administrar la evaluación. Así mismo se deberán identificar factores o variables extrañas que pueden afectar la recogida de datos en la evaluación, por ejemplo, formación de los evaluadores, o factores derivados de los instrumentos o del contexto que pueden afectar a los datos que se recogen...
3. *Selección del diseño de investigación a utilizar para evaluar el programa:* La primera decisión será elegir entre los dos macromodelos de evaluación, el modelo no lineal (evaluación formativa), o el modelo lineal (sumativa) y para ello se deberá tener en cuenta la finalidad de la evaluación (evaluación del resultado o evaluación del proceso). No obstante, en ocasiones, y siempre teniendo en cuenta los objetivos de la evaluación la utilización de ambos modelos puede complementarse. Tras la elección primaria del macromodelo se elegirá el diseño de investigación a utilizar pudiendo optarse o bien por diseños experimentales (caso único, diseños clásicos de grupos pretest-posttest con o sin grupo de control...), o bien por diseños no experimentales como la descripción naturalista mediante técnicas observacionales. Finalmente cabe plantearse si existen amenazas a la validez interna y externa, y en ese caso determinar procedimientos para el control de estos sesgos.
4. *Calendario para la evaluación:* Cuando todas estas cuestiones se han determinado se procede a preparar el calendario de la evaluación, es decir, la secuencia con la que se procederá a la recogida de los datos.

FASE 3. IMPLEMENTACIÓN Y EVALUACIÓN DEL PROGRAMA

1. *Implementación del programa y recogida de datos para la evaluación:* Cuando el programa y la metodología de evaluación han sido diseñadas, se procede a la implementación del programa y a la recogida de datos para llevar a cabo su evaluación.
2. *Análisis de los datos:* El análisis de datos dependerá de la naturaleza de los datos recogidos y del tipo de preguntas que hayan sido formuladas en relación al programa. No será lo mismo si se trabaja con datos cuantitativos o cualitativos, así como si se requieren resultados descriptivos del programa (frecuencias, porcentajes, medias, desviaciones típicas) o si se buscan nexos causales que implican análisis inferenciales (análisis de varianza univariados o multivariados, regresión múltiple...). En esta fase el evaluador se pregunta sobre el soporte en el que va a almacenar la información obtenida, sobre el tipo

de análisis que se pueden realizar con los datos que se han recogido y para responder a las preguntas a las que se pretende dar respuesta con la evaluación. El criterio principal que debe guiar el análisis a la hora de evaluar un programa es la identificación de los efectos del programa, por un lado, su eficacia (si se han cumplido los objetivos del programa), y por otro, su efectividad (si ha tenido otros efectos positivos independientes de los objetivos). No obstante, cabe destacar que el análisis de los datos no debe consistir en la aplicación mecánica y ciega de las técnicas previstas en el diseño, sino un proceso mucho más reflexivo e interactivo en el que cada resultado puede influir en las decisiones acerca de los siguientes análisis a llevar a cabo.

FASE 4. ELABORACIÓN DEL INFORME DE LA EVALUACIÓN DEL PROGRAMA

Después de implementar el programa y realizar la evaluación del mismo se procede a realizar el informe de la evaluación en el que se deben especificar todas las cuestiones relativas tanto al programa, a la metodología de evaluación, al análisis realizado con los datos y a los resultados obtenidos, con las recomendaciones pertinentes en función de los hallazgos encontrados. Un informe de evaluación debe contener la siguiente información:

1. *Portada del Informe:* Debe contener la denominación del programa evaluado, el nombre del evaluador o evaluadora que lo ha realizado, así como la fecha de redacción del informe.
2. *Resumen:* Debe incluir lo esencial de la evaluación, es decir, los objetivos de la evaluación, una breve descripción del programa y de su procedimiento de implementación, una breve descripción de la metodología y los instrumentos de evaluación utilizados, así como los resultados más relevantes obtenidos.
3. *Objetivos e hipótesis del estudio de evaluación del programa:* Se deben explicitar los objetivos de la evaluación del programa, es decir, de la investigación llevada a cabo con él, y así mismo se deben describir todas las hipótesis que se han formulado sobre los efectos del programa, o sobre otras cuestiones asociadas a la evaluación...
4. *Diseño y procedimiento de la investigación:* Se informa del diseño utilizado en la investigación llevada a cabo para evaluar los efectos del programa, así como del procedimiento seguido, es decir, las fases y tareas realizadas en el transcurso del estudio.
5. *Participantes:* Se describe de forma detallada las características de la muestra con la que se ha llevado a cabo la evaluación del programa (número de participantes y características de los mismos) y el procedimiento de selección muestral utilizado.
6. *Descripción del programa de intervención:* Se informa sobre las razones por las que se lleva a cabo el programa e interés del mismo, la fundamentación teórico-empírica sobre la que se sustenta (conceptos, teorías, estudios...), los objetivos generales y concretos del programa, las características, la configuración, el procedimiento de la implementación llevado a cabo describiendo las actividades o acciones realizadas...
7. *Metodología e instrumentos de evaluación del programa:* Se aporta información sobre la metodología de evaluación empleada, se describen los instrumentos de evaluación

utilizados (qué variables evalúan, como se han aplicado, corregido e interpretado los datos que se registran con éstos, garantías científicas de los instrumentos: datos de fiabilidad y validez...), la secuenciación llevada a cabo en la evaluación (sesiones de evaluación y temporalización de las mismas), variables extrañas controladas para garantizar la validez de los datos recogidos...

8. *Resultados:* Se explican los análisis realizados con los datos para responder a las preguntas planteadas y se exponen los resultados obtenidos de forma ordenada. Se presentan los resultados cuantitativos en tablas y gráficos, así como los resultados cualitativos complementarios.
9. *Conclusiones:* Se clarifica en qué medida se han encontrado resultados significativos que confirman la eficacia (ha cumplido los objetivos que el programa se proponía) y/o la efectividad del programa (ha tenido otros efectos positivos independientes de los objetivos). Por un lado, se sintetizan los resultados más relevantes y se aportan explicaciones a los mismos (factores o variables que han desempeñado un papel determinante en la obtención de estos resultados...), se contrastan los resultados con las hipótesis propuestas evidenciando si éstas se confirman o refutan, y también se contrastan los resultados con los obtenidos en otros estudios que o bien se han realizado con el mismo programa o bien se han llevado a cabo con programas que tenían objetivos similares. Por consiguiente, se procede a la presentación de las principales conclusiones del trabajo las cuales se debaten a la luz de parámetros teóricos y empíricos.
10. *Recomendaciones:* Teniendo en cuenta los hallazgos encontrados en el estudio de evaluación del programa, para finalizar el informe se procede a formular un conjunto de recomendaciones que pueden permitir mejorar el programa y por consiguiente sus efectos. Por un lado, se pueden identificar diversos aspectos positivos que deberían permanecer en subsiguientes implementaciones, tanto en lo referente al programa como a su evaluación, y por otro, se deben recomendar cambios que permitirían mejorar puntos débiles o bien del programa o bien de la metodología y los instrumentos de evaluación utilizados para evaluar sus efectos.

En relación a los informes de evaluación de un programa, se pueden plantear dos tipos de informes, un informe extenso que recoge todas las cuestiones que se han planteado previamente de forma detallada o un informe breve, que conteniendo básicamente los mismos puntos, éstos se describen de forma sintética, permitiendo una lectura ágil de el programa, la evaluación y los resultados de la misma.

8.3.2. CUESTIONES METODOLÓGICAS

Como ya se ha señalado previamente, la evaluación de intervenciones supone la sistemática investigación de los resultados de una intervención. Para llevar a cabo este objetivo, metodológicamente se requieren dos herramientas: (1) La determinación del diseño de investigación a emplear; y (2) Los instrumentos de medida o procedimientos de recogida de datos con los que se medirán los efectos de la intervención en las variables dependientes. La

adecuabilidad de estas dos herramientas, junto a una adecuada implementación del programa, y el correcto análisis de los datos, permitirán una validación correcta del programa.

8.3.2.1. PRINCIPALES DISEÑOS EN LA INVESTIGACIÓN EVALUATIVA

Los diseños para la evaluación de programas pueden variar desde el más riguroso experimento de campo a la descripción naturalista, sin olvidar, claro está, la diversidad de diseños cuasi-experimentales o los correlacionales.

De los diseños experimentales más empleados para evaluar el efecto de una intervención, Fernández-Ballesteros (1992g) describe 6 diseños. Algunos de estos no incluyen grupo de control siendo sólo útiles para dar respuesta a cuestiones relacionadas con el programa, mientras que otros incluyen más de un grupo, lo que permitirá comparaciones entre programas.

- *Diseño pretest-posttest con grupo de control equivalente:* Se identifican los estudiantes o grupos muestrales. Se aplica un pretest a todos antes de realizar la intervención con la finalidad de medir las variables sobre las que se hipotetiza va a tener un efecto la intervención. Posteriormente, son distribuidos al azar para desempeñar la condición experimental (grupo que recibirá X = programa a evaluar) y la de control (no recibe X). Se aplica el programa de intervención a los experimentales, mientras que los de control o bien no realizan nada o bien realizan otro tratamiento. Finalmente se aplica un posttest a todos. Las diferencias significativas pretest-posttest en las variables dependientes entre los experimentales y control se atribuyen al programa de intervención.
- *Diseño posttest con grupo de control equivalente:* La diferencia con el diseño previo es que no se opera evaluación pretest, es decir, previa a la intervención.
- *Diseño pretest-posttest con grupo de control no equivalente:* En primer lugar, se identifica a los sujetos (por ejemplo, 10 aulas preescolares de dos barrios) que van a pasar por X (tratamiento o programa, por ejemplo, programa de inmersión en la lengua), y en segundo lugar, se identifica a los sujetos que no van a recibir la intervención pero que deben ser semejantes a los anteriores, recogiendo información sobre las características de semejanza (por ejemplo, se seleccionan 5 aulas que tengan semejante nivel que los experimentales en variables como comprensión, vocabulario...). La asignación a ambas condiciones no se realiza al azar, sino por semejanza. Se aplica la evaluación pretest a ambos, y posteriormente el grupo experimental recibe X mientras que el de control o no recibe nada o recibe otro tratamiento. Finalmente se administra la evaluación posttest a todos.
- *Diseño de series temporales con grupo experimental sólo:* En primer lugar, se selecciona una medida de los resultados que puede ser recogida repetidamente. Se decide sobre la composición del grupo experimental (puede ser el mismo grupo al que se evalúa en las variables seleccionadas repetidas veces) y se recogen al menos tres mediciones tomadas regularmente antes de aplicar X (tratamiento o programa). Se implementa X, es decir, el programa y se continúa registrando medidas en los mismos intervalos de tiempo después de aplicar X.

- *Diseño de series temporales con un grupo de control no equivalente:* Se identifica el grupo que va a recibir el tratamiento o intervención y posteriormente se selecciona la variable dependiente que pueda ser registrada repetidamente. A continuación se identifica un grupo semejante al grupo experimental del que se puedan recoger las mismas mediciones. Se requiere como mínimo el registro de tres mediciones sucesivas en el tiempo (en experimentales y control) antes de administrar el programa a los experimentales. Se implementa el programa a los experimentales y se continua registrando información a iguales intervalos (por ejemplo, tres mediciones antes de la intervención y otras tres después de la misma) en las dos condiciones.
- *Diseño pre-experimental antes-después:* Se aplica una evaluación pretest al grupo experimental, posteriormente se implementa el programa, y en tercer lugar, se aplica una evaluación posttest. Este diseño puede ser útil para comprobar la efectividad de una serie de cursos de formación (variable independiente o programa) que pueden ser evaluados, por ejemplo, mediante una prueba de conocimientos (variable dependiente).

Además, de los diseños experimentales y cuasi-experimentales, en ocasiones se utilizan estudios correlacionales. Estos estudios pueden confirmar algunas relaciones hipotetizadas y, en este sentido, pueden utilizarse para corroborar hipótesis, aunque no para su contraste y su confirmación.

A pesar de que inicialmente la evaluación de programas implicaba la elección de un diseño experimental o cuasi-experimental, con la restricción que éstos comportan para la búsqueda de relaciones de causalidad debido a los controles que deben establecerse, algunos autores (Anguera, 1994a) consideran que ya no es admisible la aplicación de las pautas propias de la experimentación al ser conscientes del gran número de variables extrañas que se hallan presentes (organísmicas, del sujeto, de funcionamiento institucional, de implementación de los programas, de funcionamiento comunitario, etc.) en la totalidad de las intervenciones, y las dificultades, muchas veces insalvables, que plantea su control mediante las técnicas habituales (mantenimiento de valores constantes, aleatorización, balanceo, contrabalanceo...). Con esta perspectiva Anguera (1994a, 1995) propone la metodología observacional como método básico.

En general se entiende por diseño el conjunto de decisiones técnicas que vinculan los procedimientos de investigación a la función de dar respuesta a las cuestiones evaluativas, o en palabras de Anguera (1994a, 1995) la estructura que actúa como guía para facilitar la ordenación de los datos desde su obtención, almacenamiento y posterior tratamiento. En la selección del diseño debe tenerse en cuenta que éste sea apropiado a las características del programa y, al tipo de información necesaria para tomar decisiones sobre el mismo, para lo que requiere apoyarse en un examen escrupuloso de las alternativas disponibles. Anguera (1995) clasifica los diseños en tres grandes grupos. Una síntesis de la clasificación de esta autora se expone en el Cuadro 44.

Cuadro 44. Tipos de diseños de investigación en evaluación de programas (Anguera, 1995)

TIPOS DE DISEÑOS DE INVESTIGACIÓN	
Diseños no Experimentales	• <i>Diseños diacrónicos</i>: Mediante este diseño se consigue reunir información de un usuario del programa (o unidad mínima, como la evaluación de una intervención interactiva verbal en programas de prosocialidad o en la de una terapia familiar) a lo largo del tiempo necesario (según los casos, durante la aplicación del programa, o además, en un seguimiento posterior al tratamiento), registrando sistemáticamente a partir del sistema de categorías elaborado a lo largo de un período de tiempo. • <i>Diseños Sincrónicos</i>: Cuando se desea conocer la distribución de un grupo de usuarios respecto a la aplicación simultánea de varios sistemas de categorías, así como la evaluación en varios niveles simultáneos de respuesta, la estrategia de registro más adecuada es la de eventos clasificados de forma cruzada, de forma que no se requiere ningún tipo de continuidad entre eventos sucesivos, sino que resulta al ser aprehendidas secuencias conductuales que son clasificadas desde varias dimensiones. • <i>Diseños mixtos Diacrónicos/Sincrónicos</i>
Diseños con grado parcial de experimentalidad	• <i>Diseño con grupo de control no equivalente</i>: Este diseño engloba varias modalidades y se configura a partir de grupos a los que no se aplica el programa suponiendo que sus miembros tienen características semejantes a las de quienes integran el grupo experimental, aunque no son equivalentes. • <i>Diseño de discontinuidad en la recta de regresión</i>: Se trata de un diseño de tipo pre-postratamiento en el que el método de asignación de los sujetos a las condiciones de intervención son conocidas, aunque no aleatorias. A partir de las puntuaciones en el pretest se asignan los sujetos a la condición experimental y control. • <i>Diseños cuasiexperimentales con sujeto único</i>: (1) Diseños intraseries; (2) Diseños interseries temporales; (3) Diseños de series temporales interrumpidas; y (4) Estrategias combinadas.
Diseños experimentales	• Diseños experimentales de caso único. • Diseños clásicos de grupos.

En relación a los diseños, métodos y técnicas aplicables en la investigación evaluativa, García Ramos (1992) distingue 2 grandes enfoques, el cuantitativo-experimental y el cualitativo-naturalístico-descriptivo que representan los extremos de un continuo (Cuadro 45).

Cuadro 45. Paradigmas Cuantitativo-Cualitativo: Características y diseños (García Ramos, 1992)

	ENFOQUE CUANTITATIVO	ENFOQUE CUALITATIVO
Características definitorias	Lógico, cerrado, experimentalista, conductual, mecanicista, explicativo, hipotético-deductivo, racional, objetivo.	Fenomenológico, abierto, naturalístico, dinámico, descriptivo, inductivo, subjetivo.
Orientación	Al producto	Al proceso

Objetivos metodológicos	Identificación de relaciones causales.	Descripción general del programa y desde la perspectiva metodológica.
Metodología-diseños	Diseños complejos y sofisticados: experimentales, cuasiexperimentales, descriptivos, longitudinales... Preocupación por garantizar la validez interna (control) y externa (generalización). Usa el control y la comparación de grupos, técnicas de muestreo, manipulación activa de variables independientes y técnicas estadísticas de análisis de los datos	Diseños simples Métodos etnográficos, estudio de casos, observación participante, análisis de interacciones, estudios antropológicos... La evaluación es observación e interpretación, comprensión de lo que sucede.
Metodología-instrumentos	Fiabiles y validos: tests, pruebas objetivas, cuestionarios...	Cuestionarios, estudio de casos, observación, anecdóticos, entrevistas...
Variables	Definidas y operacionalizadas previamente.	Algunas definidas previamente y otras emergentes en el curso de la evaluación
Evaluador	Se sitúa como persona independiente al fenómeno a evaluar.	Interacciona con el fenómeno estudiado con la finalidad de comprender las relaciones que se dan dentro del programa, entre los participantes, desde dentro mismo del programa.

La evaluación de programas comporta en la práctica la aparición de un buen número de problemas, en gran parte de carácter metodológico, que no siempre se han resuelto, observándose lagunas especialmente en relación al procedimiento. De aquí que una de las primeras preocupaciones que plantea la evaluación de programas debe ser la elaboración de un adecuado diseño.

En la actualidad, la metodología cualitativa está recobrando terreno en el área de las ciencias del comportamiento, y volviendo a ser considerada como una metodología importante en el ámbito de la evaluación psicológica, dada su relevancia en la manera de entender a los seres humanos y a la naturaleza de sus interacciones entre ellos y con su entorno. Por el auge en la utilización de esta metodología, Anguera (2008) ha realizado una revisión teórica de esta perspectiva metodológica mostrando sus bondades. Para ello, hace un recorrido a través de sus principales características y exigencias metodológicas, así como de sus fases a cubrir.

Finalmente, esta metodóloga aboga por una postura conciliadora y complementaria entre las vertientes cuantitativa y cualitativa de la metodología. Mientras que tradicionalmente las dos grandes vertientes metodológicas han mantenido una pugna que dicotomizaba a sus seguidores, en los últimos años, por fortuna, se ha ido instaurando la cordura metodológica necesaria para iniciar una vía fructífera de complementariedad en evaluación de programas. Ha resultado un proceso duro y difícil, no exento de incomprensiones desde ambas orillas respecto a intentos iniciales de colaboración, pero que en la actualidad se halla en un proceso imparable que resulta sumamente fértil y prometedor.

El nudo gordiano de la cuestión se hallaba, por un lado, en que por parte de los seguidores de la *tradicional metodología cualitativa*, una vez se había recogido la información, sea en forma de notas de campo, diarios, transcripción de entrevistas en profundidad, etc.,

material valorado como de gran riqueza informativa, todo el esfuerzo se dedicaba unidireccionalmente a 'enlazar' datos, elaborar memorandos (comentarios reflexivos), analizar el contenido contando frecuencias, generar mapas conceptuales, organizar matrices o redes para condensar la información recogida, proponer nuevas teorías a partir de las relaciones detectadas por inspección visual y la posterior reflexión... pero siempre había una carencia metodológica latente en todas estas actividades, y era la subjetividad, o, la carencia de objetividad demostrada, debilidad que ni siquiera se situaba en el epicentro de las preocupaciones de los que realizaban investigaciones cualitativas, mucho más preocupados habitualmente por plasmar sus discursos reflexivos o por construir teoría (Anguera, 2008).

Pero, por otro lado, los seguidores de la tradicional metodología cuantitativa operaban exclusivamente con valores numéricos procedentes de medida mediante instrumentos estándar que cumplían los requisitos métricos, o de constatación del orden, permitiendo descubrir, verificar o identificar relaciones simétricas o no entre conceptos que derivaban de un esquema teórico elaborado de acuerdo con los criterios que regían cada una de las situaciones que interesaba estudiar, exigiendo, para llevar a cabo el contraste de la hipótesis, los requisitos de representatividad y aleatorización, lo cual comportaba unas adecuadas técnicas de muestreo, y después se llevaban a cabo técnicas cuantitativas de análisis que podían alcanzar elevada sofisticación; no obstante, siempre había una carencia metodológica latente en todas estas actividades, que era la omisión voluntaria de cualquier información procedente de la descripción de la realidad, fuesen episodios, documentos personales, registros descriptivos, autoinformes... a los cuales se los consideraba como 'material de chatarra' para el estudio del comportamiento, y, por consiguiente, se despreciaba la riqueza de información que podían aportar (Anguera, 2008).

Por supuesto, los planteamientos son divergentes, y mucho se ha discutido acerca de si comportan un diferente cuño epistemológico, así como de la necesaria elección del paradigma. Pero la cuestión que interesa es hasta qué punto es factible en evaluación de programas aunar sus ventajas o puntos fuertes a la vez que se compensen sus carencias o puntos débiles. La tarea no es fácil, pero si es factible, y la clave se reduce a dos cuestiones fundamentales: 1) Por una parte, cubrir las respectivas deficiencias de ambos colectivos de investigadores, considerando la extraordinaria riqueza que aportan los registros de conducta, la conducta verbal transformable a material documental, y los materiales documentales, con lo cual alcanzarán el rango de 'datos', a la vez que se fortalecen todas las cautelas y acciones metodológicas tendentes a controlar la calidad del dato con el fin de garantizar la necesaria objetividad propia de cualquier estudio científico; y 2) Por otra parte, aprovechar la pertinencia de los instrumentos no estándar, que, mediante los sistemas de categorías (en los diseños unidimensionales) y los formatos de campo (en los diseños multidimensionales), así como la combinación entre ambos (en los diseños multidimensionales), dotan a los investigadores de importantes herramientas que posibilitan la transformación de datos cualitativos a datos tratables analíticamente de forma cuantitativa.

En este sentido, la metodología observacional, que da soporte básico a la evaluación de programas de baja intervención, ha sido pionera en lograr esta complementariedad desde su origen –considerado tradicionalmente como una metodología cualitativa- hasta los desarrollos analíticos posteriores, que prácticamente no presentan ninguna limitación respecto a los tradicionales estudios cuantitativos.

Con estas nuevas consideraciones metodológicas, se supera la crítica de la endémica debilidad metodológica de los enfoques cualitativos en evaluación de programas, a la vez que se consiguen importantes avances consistentes en el uso de recursos metodológicos sofisticados que permiten un rigor mucho más elevado. Una idea básica y esencial que forma parte de esta nueva concepción de la metodología cualitativa en su complementariedad con la cuantitativa consiste en que, en función del diseño planteado y de la naturaleza de los datos, procederá una u otra técnica analítica, teniendo presente que habitualmente se dispone de datos categóricos. En cualquier caso, si la metodología cualitativa nos ayudó en la obtención de datos que aportan una gran riqueza informativa, la cuantitativa nos suministra los recursos para su análisis más conveniente. La lógica sucesión de etapas de forma organizada debe permitir este cambio de perspectiva (de *lo* cualitativo a *lo* cuantitativo) de forma pacífica, sin tensión interna en el seno del procedimiento a seguir. Ambas vertientes metodológicas pueden beneficiarse mutuamente entre sí, y son muchas las ocasiones en que se utilizan de manera conjunta, dando garantía de su complementariedad. Se abre paso, con ello, a los diseños de evaluaciones multimétodo, que pueden seguir una estrategia combinada o integrada, siempre planificada acorde a los objetivos del programa y del tipo de evaluación elegida (Anguera, 2008).

Como enfatiza Anguera (2008), es importante señalar el papel que han desempeñado las nuevas tecnologías y los avances informáticos en esta tarea de acercamiento entre las posiciones metodológicas cualitativa y cuantitativa, si bien es cierto que existen programas que siguen los planteamientos conceptuales más radicales. Es muy amplio el elenco de programas de que se dispone, así como la inusitada velocidad con que aparece nuevo software en el mercado y en la comunidad científica. En la actualidad ya no nos sorprende ni que se realicen determinados análisis estadísticos a partir de textos, ni que determinados programas, como el Hyperbase, puedan llevar a cabo análisis de conglomerados o análisis factorial, o que otros que tradicionalmente se han aplicado a registros sistemáticos de conducta, como el Thème, puedan extraer T-Patterns a partir de la transcripción de sesiones de psicoterapia, sea tradicional o tecnológicamente mediada.

En el ámbito de la evaluación de tratamientos clínicos, Haynes y Williams (2003), subrayan que los tratamientos estandarizados pueden no ser completamente efectivos en el caso de clientes con múltiples problemas de conducta, que se afectan entre sí y que están determinados por variables causales interrelacionadas. Los mecanismos de acción de los tratamientos estandarizados puede que no siempre consideren las variables causales más importantes de los problemas de conducta de un cliente. Por ello sugieren que el análisis funcional de la conducta posibilita en estos casos el diseño de tratamientos individualizados, mediante la evaluación integrada de los problemas conductuales del cliente, de las variables causales importantes, y de las relaciones funcionales entre las mismas. Los parámetros de un análisis funcional pueden explicitarse mediante Modelos Funcionales Analíticos de Caso Clínico (FACCM *Functional Analytic Clinical Case Model*). Se trata del modelo gráfico vectorial de un análisis funcional, cuyo objetivo último es facilitar la toma de decisiones sobre el tratamiento, ayudando al clínico a estimar la magnitud del efecto de cada variable causal en los problemas conductuales del cliente, de forma tal que puedan seleccionarse los tratamientos más efectivos.

Desde esta perspectiva, Haynes, Leisen, y Elaine (1997) consideran que el diseño de un programa de tratamiento individualizado en terapia conductual en buena medida depende de los datos pretratamiento obtenidos como parte de una evaluación multimétodo multimodal. El

análisis funcional es la integración de diversos elementos para la toma de decisión de un tratamiento individualizado: a) la importancia relativa, las interrelaciones, y las consecuencias de los problemas de conducta del paciente; b) los objetivos del tratamiento; y c) la relativa modificabilidad, interrelaciones y peso de las variables causales. El análisis funcional puede representarse visualmente mediante el Modelo Funcional Analítico de Caso Clínico (FACCM *Functional Analytic Clinical Case Model*), una representación gráfica vectorial de las variables y sus relaciones funcionales.

8.3.2.2. PROCEDIMIENTOS E INSTRUMENTOS PARA EVALUAR EL CAMBIO

8.3.2.2.1. Características básicas de los instrumentos de evaluación

De la Orden (1990) propone una guía para tomar decisiones sobre instrumentación en evaluación destacando los siguientes puntos: 1) Especificación de objetivos; 2) Sensibilidad a los efectos del programa; 3) Factores que afectan a los resultados de las medidas; 4) Sesgos de la medida y sus efectos en la evaluación; 5) Pruebas sobre nivel y fuera de nivel; 6) Tests comerciales versus tests “*ad hoc*”. En opinión de este autor para medir los efectos de un programa, el proceso evaluador deberá especificar las variables que se intentan medir, seleccionar y/o elaborar instrumentos pertinentes y, juzgar su adecuación para medir tales variables.

Para Fernández-Ballesteros (1992) las características principales de los instrumentos de evaluación de intervenciones son cuatro: 1) *Objetividad*: Que los procedimientos de recogida de información estén libres de los errores procedentes de la interpretación del evaluador; 2) *Estabilidad*: Que la medida no presente excesivas fluctuaciones debidas a su propia naturaleza, de forma que al juzgar los cambios debidos al programa éstos no puedan confundirse con fluctuaciones en el tiempo del propio instrumento; 3) *Exactitud, Precisión o Validez*: Que la medida elegida refleje lo que se pretende juzgar, es decir, que las técnicas utilizadas sean una medida, lo más precisa posible, de lo que se pretende dar cuenta; y 4) *Sensibilidad al cambio*: Que el procedimiento elegido sea sensible ante el programa de acción propuesto.

Básicamente, los instrumentos que se utilizan para evaluar una intervención, un programa o un tratamiento, no difieren de los empleados en evaluación psicológica en general, ya que si queremos evaluar el impacto de una intervención en la inteligencia utilizaremos típicos procedimientos de medida como los tests de inteligencia, aptitudes y de desarrollo intelectual; mientras que si el objetivo es disminuir el fracaso escolar, uno de los mejores indicadores será el número de suspensos o los resultados en pruebas de rendimiento académico.

8.3.2.2.2. Utilización de múltiples técnicas: Instrumentos cualitativos y cuantitativos

Al realizar un breve recorrido histórico, se puede observar que en sus inicios, en la evaluación de programas se hacía uso casi exclusivo de la metodología cuantitativa, guiada por teóricos experimentalistas como Campbell y Scriven. Con ello se aportaban conocimientos acumulativos, donde la evaluación sumativa ocupaba un lugar relevante. En la década de los 70 aparecen nuevos planteamientos de investigadores como Patton que abogan por una metodología cualitativa. En la actualidad, muchos autores consideran que la incompatibilidad de

ambas posturas es falsa, y que la elección de una u otra debe depender del propio programa, pudiéndose considerar ambas como complementarias.

Los métodos cualitativos en opinión de Patton (1987) permiten un estudio más profundo y detallado cuando se trata de unos pocos casos, siendo sus características principales: 1) El evaluador se centra en la observación del desarrollo del programa, informa de las posibles variaciones y explora las diferencias individuales; 2) Los tipos de análisis emergen de la propia observación y registros, no están guiados por hipótesis previamente formuladas; 3) Adquiere una enorme importancia la empatía; 4) Se asume que la totalidad es más que la suma de partes; y 5) Se concibe el programa como algo dinámico.

Hernández (1995) en el contexto del debate cuantitativo-cualitativo destaca diferencias en los objetivos de ambos métodos, que se exponen en el Cuadro 46.

Cuadro 46. Objetivos de los métodos cuantitativos y cualitativos (Hernández, 1995)

MÉTODOS CUANTITATIVOS	MÉTODOS CUALITATIVOS
• Búsqueda de objetividad o de intersubjetividad. • Establecimiento de generalidades. • Obtención de datos. • Sujeto de evaluación: individuo en poblaciones. • Diseño planificado. • Dirigidos a la evaluación de programas de largo alcance. • Evaluación del resultado.	• Búsqueda de subjetividad de implicados representativos. • Conocimiento de particularidades. • Obtención de discursos. • Sujeto de evaluación: individuo en grupos. • Diseño emergente. • Dirigidos a evaluaciones de programas locales. • Evaluación del proceso.

En relación a los instrumentos de evaluación, como puede observarse en el cuadro 47, los instrumentos empleados en la evaluación de programas o intervenciones son semejantes a los utilizados para la evaluación psicológica de un individuo (entrevistas, observación, autoinformes, tests objetivos, técnicas proyectivas...), y además, se incorporan otras herramientas (técnicas etnográficas, encuestas, archivos, indicadores estadísticos...). Al abordar las alternativas metodológicas existentes resulta adecuado plantear el empleo de instrumentos cualitativos y cuantitativos ya que su combinación puede enriquecer la evaluación del cambio producido por un programa. En el Cuadro 47 se detallan las técnicas e instrumentos empleados con mayor frecuencia.

Cuadro 47. Instrumentos de evaluación de programas, intervenciones o tratamientos

CUALITATIVOS	CUANTITATIVOS
• Entrevista abierta. • Observación no estructurada: diarios.... • Técnicas sociométricas. • Grupos de discusión. • Cuestionarios o encuestas abiertas. • Técnicas etnográficas: entrevista callejera... • Autobiografía • Análisis históricos: a través de distintos tipos de datos como periódicos o informes con los que se pretende establecer la historia pasada de un evento. • Método Delphi: expertos a los que se envían los	• Entrevista estructurada. • Autoinformes: el sujeto se valora a sí mismo • Cuestionarios: sobre el nivel de satisfacción del programa, o sobre la percepción subjetiva del cambio por efecto del programa... • Indicadores estadísticos y sociales: número de toxicómanos, tasas de empleo... • Observación estructurada y/o sistemática: frecuencia de conductas observables... • Archivos: calificaciones, hojas de servicio... • Medidas de rastro y erosión: número de cigarrillos,

cuestionarios y los resultados de los mismos...	botellas... • Escalas de apreciación: jueces (padres o profesores) califican al sujeto en una escala de estimación en relación a alguna conducta... • Pruebas de ejecución y de inteligencia: piezas insertadas, número de aciertos, tiempo de ejecución... • Pruebas de rendimiento: exámenes... • Cuestionarios: medidas de autoinforme sobre habilidades, nivel de satisfacción, o logros obtenidos con el programa... • Entrevistas estructuradas. •

Por otra parte, algunos autores han utilizado otros ejes de clasificación de los instrumentos. Anguera (1994a) plantea una clasificación de los instrumentos en función de su grado de estandarización, proponiendo el esquema que se presenta en el Cuadro 48.

Cuadro 48. Clasificación de los instrumentos en función de su grado de estandarización

GRADO DE ESTANDARIZACIÓN	INSTRUMENTOS	
Ausencia de Instrumento estándar	Directamente perceptible	Registro de Conducta
	No directamente perceptible	Información documental
Semiestandarización	Cuestionario, Escala, Autoinforme... Entrevista	
Instrumento estandarizado	Pruebas psicológicas psicométricas clásicas Registros psicofisiológicos	

Así, para las respuestas directamente perceptibles Anguera (1991, 1994a) sugiere el registro de conducta para lo que se aplicarán estrategias propias de la observación directa como técnica, aunque no excluye el uso de evaluaciones observacionales a partir de protocolos de registro. Los principales problemas que se plantean se refieren a la fijación de las unidades de conducta, su sistematización y categorización, y la optimización del registro, así como la profundización en el análisis de datos.

Las respuestas no directamente perceptibles por el usuario han estado cuestionadas por su riesgo de subjetivismo y por su tradicional falta de sistematización. Para estas respuestas se sugiere el uso de estrategias propias de la observación indirecta como técnica. Esta implica la existencia de conductas encubiertas que requerirán una inferencia, y, por lo tanto, una carga interpretativa que puede redundar en detrimento y menoscabo de la objetividad requerida en todo estudio científico y, en consecuencia, en la evaluación. Sin embargo, en los últimos años se va avanzando en este sentido, y el estudio de los procesos cognitivos superiores ha impulsado a empujar los límites que circundan a la observación indirecta (Anguera, 1994a).

Dentro de los instrumentos de carácter semi-estandarizado, en esta clasificación se incluyen diversas técnicas (entrevista, checklist, cuestionario...), y en cuanto al uso de instrumentos estándar existe un amplio abanico de posibilidades en función de las diferentes áreas de especialización. El aspecto más destacable en relación a los instrumentos estandarizados lo constituye la exigencia de cumplimiento de las garantías psicométricas

(fiabilidad, validez, sensibilidad, pautas de interpretabilidad y normas para distintas poblaciones...).

En un trabajo reciente sobre la evaluación de programas en contextos educativos (Pérez Juste, 2006) se sintetizan diversas técnicas que pueden utilizarse para evaluar un programa:

1. *Técnicas observacionales*: para recoger información sobre comportamientos que ocurren de forma natural o son provocados por el observador.
2. *Técnicas de análisis*: análisis de documentos como actas, escritos, expedientes académicos, filmaciones, registros...; para comprobar si algo ocurre o no, con qué frecuencia, intensidad o duración...
3. *Técnicas de encuesta*: una encuesta es un cuestionario o inventario con preguntas cerradas que pueden ser procesadas por una lectora óptica o abiertas para realizar análisis de contenido, que permite recoger información a amplios grupos de población, pero puede tener como limitación la falta de sinceridad en las respuestas del evaluado
4. *Técnicas de consulta*: se consulta a expertos que son personas que por su experiencia, formación, dedicación... pueden aportar información, ideas, propuestas, sugerencias... de interés.
5. *Técnicas de prueba*: cuando el objetivo de un programa es adquirir un conocimiento o realizar un aprendizaje (leer, nadar, hacer esquemas, redactar...) la metodología de evaluación adecuada no será recoger las percepciones subjetivas de los destinatarios sino que se debe basar en una prueba objetiva que permita valorar la adquisición de ese aprendizaje.

Este autor en la misma dirección que otros investigadores sugiere la utilizar complementaria de técnicas objetivas y subjetivas en la evaluación de un programa. Aunque desde algunos modelos la utilización de diversos tipos de técnicas no se considera adecuada, sin embargo, dependiendo de los objetivos que tenga la evaluación del programa, y dependiendo de las características del propio programa se pueden utilizar de forma complementaria técnicas cuantitativas y cualitativas, técnicas objetivas y subjetivas, ya que la información que aportan las diversas fuentes puede enriquecer la comprensión del programa objeto de evaluación.

8.3.2.3. PROBLEMAS METODOLÓGICOS RELACIONADOS CON EL PROGRAMA

Para Anguera (1994a) un programa puede tener problemas en tres niveles, defectos de planteamiento del programa, irregularidades en su implementación, y falta de monitorización. Los *defectos en el planteamiento* de un programa pueden ser debidos a distintas razones:

- Programas mal definidos que requieren en consecuencia una posterior elaboración para ser implementados.
- Vaguedad en los objetivos, lo cual no permite una suficiente base de sustentación para establecer las directrices del programa.
- Fines no realistas o ilusorios, los cuales impedirán un adecuado establecimiento de acciones, ya que su seguimiento sería baldío.
- Objetivos contrapuestos, que generarían una contradicción interna en la estructura de programa.

- Falta de claridad en el enunciado de las tareas requeridas.
- Inexistencia de organigrama relativo a la secuenciación o concurrencia de las distintas acciones de que se compone el programa.
- Falta de calendario o previsión temporal.
- Imprecisión respecto al estudio empírico, lo cual se diversifica a su vez en uso y validación de instrumentos, naturaleza de los datos y forma de obtención, unicidad de criterios en el almacenamiento del material de archivo, indicadores a utilizar, perspectivas de análisis de datos....

El esfuerzo por resolver estos problemas redundará en un aumento de la factibilidad evaluativa del programa, lo que no quiere decir, por supuesto, que con ello mejore la valoración desde todos los tipos de validez, pues si bien la interna se eleva, pueden en cambio permanecer inamovibles otros tipos, como la externa (si los resultados no son generalizables) o la social (si la elaboración del programa no estuvo precedida de un análisis previo de necesidades que hiciera a este pertinente).

Por otro lado, pueden surgir problemas derivados de *irregularidades en la implementación* del programa de dos tipos: 1) Falta de formación y/o monitorización en las personas encargadas de llevarlo a cabo; y 2) Errores asistemáticos de naturaleza aleatoria y, en consecuencia, no previsibles, por ejemplo, rechazo de sujetos más allá de los criterios establecidos, incorrecta cumplimentación de protocolos... Finalmente, un tercer problema en la implementación del programa es *la falta de monitorización*. Como señala Anguera, no es todavía frecuente institucionalmente el sometimiento del equipo de profesionales que trabajan en la implementación de un programa, a un asesoramiento que mejore la implementación, y con ello, en consecuencia, a evitar la desazón que produce la recogida de datos con importante volumen de “missings” o la falta de previsión para evitar la mortalidad experimental, o la duda sobre cómo analizar previamente la validez de un determinado instrumento, o las dificultades para efectuar un análisis de costo-beneficio.

8.3.2.4. EL ANÁLISIS DE LOS DATOS EN LA EVALUACIÓN DE LOS EFECTOS DEL PROGRAMA

La calidad del análisis depende, aparte de la competencia e independencia del analista, de la calidad del diseño evaluativo y de su ejecución. En el proceso de análisis, se elige la técnica adecuada en función del diseño, las variables y sus medidas, se aplica y se interpretan los resultados. Existen procedimientos estadísticos como el ANOVA, el ANCOVA, o la regresión múltiple que son empleados con frecuencia cuando se realiza un estudio experimental de validación del programa.

Con la finalidad de mejorar el proceso de análisis estadístico de datos se pueden seguir las siguientes pautas:

- *Análisis exploratorio de los datos*, sometidos los resultados a los procedimientos rutinarios de descripción: frecuencias, estadísticos univariados como media, mediana, moda, desviación típica...
- *Análisis inferencial primario* que permite tomar decisiones muy ajustadas respecto al tipo de análisis que se debe aplicar, es decir, se trata de decidir el modo de analizar formalmente los datos (ANOVA, ANCOVA, regresión múltiple...). Se procede al análisis de una variable

cada vez, que coincidirá con un objetivo del programa, lo que permitirá determinar cuantos objetivos se han logrado. Cuando se trabaja con múltiples variables dependientes se tiende a utilizar diseños multivariados (MANOVA).

- *Reanálisis* comparando diversos procedimientos.
- *Interpretación* que implica la comparación de resultados entre los diversos análisis, determinación del significado y coherencia interna de los resultados (tamaño del efecto) y del significado, y coherencia externa.

En la misma dirección, Botella y Barriopedro (1995) ofrecen una guía orientativa de los pasos que podrían seguirse en el análisis de datos, enumerando ocho pasos:

- Establecer el sistema de codificación de las variables cualitativas y ordinales y la métrica de las cuantitativas (obtener la fiabilidad intercodificadores con submuestras aleatorias de los datos).
- Preparar la matriz de unidades por variables.
- Obtener un primer análisis descriptivo de cada variable y de combinaciones seleccionadas de dos variables, obteniendo una primera versión de los instrumentos descriptivos. Implica: Detectar valores o combinaciones de valores imposibles; Identificar valores atípicos y decidir sobre ellos; Recodificar mediante agregados o subdivisiones de categorías o crear nuevas variables cuantitativas mediante la suma o ponderación de otras variables; Volver a obtener los instrumentos descriptivos tras las eventuales modificaciones derivadas de los puntos anteriores.
- Analizar la presencia y aleatoriedad de la distribución de valores ausentes mediante las técnicas apropiadas.
- Aplicar técnicas inferenciales para determinar la significación de los efectos aparentes en los datos.
- Estimar la magnitud del impacto en los distintos indicadores.
- Aplicar las técnicas inferenciales para decidir sobre la consecución de las magnitudes criterios establecidas como objetivo del programa.
- Aplicar técnicas meta-analíticas, cuando proceda, para mostrar y comparar tamaños del efecto diferencial.

En relación a algunos principios que deben presidir el proceso de análisis, Botella y Barriopedro (1995) enfatizan que el análisis de los datos no debe consistir en la aplicación mecánica y ciega de las técnicas previstas en el diseño, sino un proceso mucho más reflexivo e interactivo en el que cada resultado puede influir en las decisiones acerca de los siguientes. Además, los autores subrayan que en caso de duda entre dos técnicas o procedimientos de análisis es mejor hacerlo con ambos. Y finalmente, sugieren que si bien es conveniente explotar los datos al máximo, se debe realizar una buena y clara descripción de lo principal.

8.3.3. ÁMBITOS DE APLICACIÓN DE LA EVALUACIÓN DE PROGRAMAS EN ESPAÑA

A modo de conclusión del capítulo, se procede a realizar una revisión de las aportaciones de psicólogos españoles a la evaluación de intervenciones, programas o tratamientos en España en las últimas décadas, constatando que muchas de las aplicaciones en evaluación de programas, especialmente en los primeros años, se han operativizado en el campo educativo,

contexto en el que se han formulado una serie de condiciones para que estas valoraciones sean aceptables. No obstante, la revisión llevada a cabo evidencia que en los últimos años la evaluación de programas ha sido especialmente relevante en el ámbito de los servicios sociales, en el ámbito educativo, así como en el clínico y de la salud. En concreto se han llevado a cabo evaluaciones de programas en diversos contextos con variados objetivos tales como:

- **Programas en servicios sociales de protección de menores** (Arruabarrena, 2009; Arruabarrena, Berrondo y García, 1994; Arruabarrena y De Paul Ochotorena, 2002; De Paul Ochotorena, 1993; Fernández-Ballesteros y Fernández del Valle, 1991; Fernández del Valle, 1992, 1994, 1995; Fernández del Valle, Álvarez, y Bravo, 2003; Fernández del Valle y Fernández Ballesteros, 1993; Fernández del Valle y García-García, 1994).
- **Programas destinados a la tercera edad**, intervenciones llevadas a cabo para incrementar la actividad social en ancianos (Fernández del Valle, Errasti, y Pérez, 1995) sus habilidades interpersonales (Pelechano, 1995; Pelechano y De Miguel, 1992, 1993), su competencia espacial (Calero y García, 1995), sus aptitudes físicas (Díaz, Díaz, Acuña, Donoso, Nowogrodsky, 2002; Godoy et al., 1995), la memoria (Montejo, Reinoso, de Andrés, Claver, 2001) o de planificación y evaluación de programas en contextos gerontológicos (Fernández-Ballesteros, García-Merita, Hernández, 2000).
- **Programas de intervención psicológica en diversas variables del desarrollo de la personalidad infanto-juvenil en contextos educativos y clínicos**, para fomentar variables tales como autoconcepto, extraversión, creatividad, cooperación, conducta prosocial, inteligencia, o inhibir la conducta disruptiva... (Esteve, 1995; Garaigordobil, 1992ab, 1994c, 1995ace, 1996ab, 1999ab, 2000, 2001, 2002, 2004ab, 2005abc, 2006, 2008ab, 2009, Garaigordobil y Berruero, 2007; Garaigordobil y Echebarria, 1995; Garaigordobil, Maganto y Etxeberria, 1996; Garaigordobil, Maganto, Pérez, y Sansinenea, 2009; Garaigordobil y Pérez, 2004; Gavino, Capafóns y Sosa, 1985; González-Portal, 1992, 1994; Hernández et al., 1994; Maganto, del Rio y Roiz, 2002; Polaino y Lizasoain, 1994; Prieto y Pérez Sánchez, 1993; Sosa, Capafóns, Gavino y Carrio, 1984).
- **Programas implementados en contextos educativos**. Entre los programas que han sido aplicados y evaluados recientemente cabe destacar programas de educación sexual (Lameiras, Carrera, Núñez, y Rodríguez Castro, 2006), desarrollo de la creatividad (Prieto, López, Bermejo, Renzulli, y Castejón, 2002), de prevención de la violencia entre iguales (Cerezo, 2006), para la prevención de abusos sexuales en la infancia (Campos y López, 2006), de desarrollo emocional y moral (Alonso-Gancedo e Iriarte, 2005, Aciego de Mendoza, Domínguez, Hernández-Hernández, 2003), o para el desarrollo de la autoestima (Cava y Musitu, 2000).
- **Programas de intervención con niños hospitalizados por cancer** (Pozo y Polaino, 1993).
- **Programas para la prevención del consumo de drogas** (Maciá, Olivares y Méndez, 1994; Santacreu, 1994; Santacreu y Frojan, 1994) **y para el tratamiento de las drogodependencias** (Fernández-Hermida y Secades, 2000; Maya y Holgado, 2007)
- **Programas de tratamiento de trastornos clínicos**, por ejemplo, en trastornos fóbicos (Capafóns, Sosa, Viña, López y Avero, 1994; Capafóns, Sosa, y Avero, 1996; Echeburúa, Corral, García y Borda, 1991, 1993; 1996d), en hipocondría (Martínez Narváez y Botella,

1995), en trastorno obsesivo-compulsivo (Sánchez Arribas, 1996), en el trastorno por angustia (Botella, 1996), en anorexia y bulimia (Calvo, 1995), en problemas sexuales (Capafóns, Sosa y López, 1986) o con el lenguaje (Angulo, 1977), en el tratamiento de la esquizofrenia (Cirici, 2002; Lemos et al., 2004) y otros trastornos mentales (Besora, Martorell y Clusa, 2000; Macía y Méndez, 1991; Muñoz y Panadero, 2002).

- **Programas de intervención en contextos penitenciarios.** Programas para el desarrollo de la toma de perspectiva en el área de las habilidades interpersonales (Beleña y Bagená, 1992; Beleña y Villarroja, 1993; Redondo, 1993), sobre las necesidades de evaluación de la intervención penitenciaria (Anguera y Redondo, 1991) o sobre instrumentos para la evaluación de programas en menores delincuentes institucionalizados (Silva, López, Brengelmann y Benedicto, 1989), tratamiento de agresores sexuales (Redondo y Garrido, 2008; Martínez, Pérez, Navarro, y Redondo, 2007)...
- **Programas de intervención psicológica con mujeres violadas** (Corral, Echeburúa, Sarasua y Zubizarreta, 1995; Echeburúa, Corral, Sarasua y Zubizarreta, 1996) **y maltratadas** (Corral, 1996; Matud, Gutierrez, y Padilla, 2004).
- **Programas de intervención social**, por ejemplo, programas dirigidos a inmigrantes y otras minorías étnicas (Hernández, Reboloso, Fernández, Pozo, y Alonso, 2000).
- **Programas de intervención aplicados en situaciones de desastre** (Salazar, Caballo y González, 2007).
- **Programas de la rehabilitación cognitiva de personas con daño cerebral** (Moreno, Mora y Saldaña, 2007).

8.4. ALGUNAS PREDICCIONES FUTURAS SOBRE LA EVALUACIÓN PSICOLÓGICA Y EL ROL DEL PSICÓLOGO EN ESPAÑA

Como se ha puesto de relieve en los capítulos previos, la evaluación psicológica en España sigue siendo tarea fundamental dentro de la actividad del psicólogo, sobre todo en los ámbitos de aplicación clásicos y en ellos la orientación teórica tendencialmente dominante es la cognitivo-conductual. Todos los cambios habidos tanto a nivel social como en otros campos están teniendo y tendrán un reflejo en el área de la evaluación psicológica. Teniendo en cuenta estas transformaciones, y en una dirección similar a la planteada por Calero (2007), cabría formular varias predicciones de futuro sobre la evaluación psicológica y sobre el rol del psicólogo en España, tales como:

- **Incremento de la cooperación interdisciplinar con otros profesionales y especialización del psicólogo en campos concretos:** Del mismo modo que está sucediendo en otros campos del conocimiento, el avance pasa por la cooperación interdisciplinar llevando a cabo un trabajo cooperativo con otros profesionales de otras disciplinas (médicos, asistentes sociales, profesores, técnicos informáticos...), lo que implica para el rol del psicólogo una dirección de especialización en un campo específico, superando la concepción del psicólogo generalista y ahondando en la dirección de profundización y especialización en un área de conocimiento específica.
- **Mayor incorporación de los psicólogos en los servicios de salud:** En España, a diferencia de lo que sucede en otros países europeos y en Estados Unidos, la presencia del psicólogo en el sistema de salud, en los servicios de salud, es bastante escasa, y esto

también se refleja en el ámbito de la evaluación psicológica. Por ello, es esperable que en futuro aumente más la presencia del psicólogo tanto en lo que se refiere a la ejecución de funciones de evaluación como de tratamiento.

- **Mayor conexión entre evaluación y tratamiento:** Aunque la evaluación conductual desde sus primeros desarrollos enfatizó la importancia de la conexión entre evaluación y tratamiento, en la actualidad, en muchas ocasiones se realiza una evaluación, no habiendo conexión entre ésta y el tratamiento subsiguiente. Hoy en día todos los psicólogos reconocen que la evaluación tiene sentido únicamente si se vincula a la intervención o al tratamiento, pero con mucha frecuencia pese a que la evaluación ponga de relieve la necesidad de tratamiento, éste no se lleva a cabo como parte del proceso.
- **Incremento de tecnología evaluativa que disponga de garantías científicas:** Otra dirección de futuro apuntará al desarrollo de estudios que aporten más garantías científicas a los instrumentos existentes, así como el desarrollo de nuevas herramientas con garantías que eviten niveles altos de inferencia en torno a los resultados.
- **Mayor sensibilidad del psicólogo hacia el código deontológico tanto en lo que se refiere a la evaluación psicológica como al resto de las actuaciones profesionales:** En la actualidad se producen diversos tipos de vulneraciones del código deontológico sin que exista un sentimiento de responsabilidad al respecto, en el futuro el psicólogo se puede predecir que aumentará la sensibilidad de los psicólogos en relación a la conducta ética y una mayor tendencia a asumir las consecuencias de sus acciones, tomando conciencia de la responsabilidad cuando éstas son negativas.
- **Desarrollo de una mayor formación multicultural así como de instrumentos que eviten sesgos relacionados con la cultura:** Los procesos migratorios masivos acontecidos en las últimas décadas en el mundo obligan a los psicólogos a adquirir una mayor formación multicultural, y esto conlleva la construcción de instrumentos de evaluación que eviten los sesgos culturales que en la actualidad tienen muchos instrumentos de evaluación.

Como han subrayado Kaplan y Saccuzzo (2006), el futuro de las pruebas psicológicas depende de muchos aspectos y desarrollos. Los aspectos profesionales incluyen asuntos teóricos, como la utilidad del concepto de rasgo en oposición al índice de ajuste; la eficacia de las pruebas y la predicción actuarial contra la clínica. Los aspectos morales incluyen derechos humanos, como el derecho a no ser evaluado, el derecho a no ser etiquetado, y el derecho a la privacidad. Otro aspecto ético que se relaciona con los usuarios y desarrolladores de pruebas es la lealtad dividida que puede resultar de administrar una prueba a un individuo para una institución: ¿los derechos de quien son primero? Asimismo, los profesionales tienen el deber ético de proporcionar y entender la información necesaria para usar una prueba de forma apropiada. Finalmente, los aspectos sociales como la deshumanización, la utilidad de las pruebas y el acceso a los servicios de pruebas, también conforman el campo actual de las pruebas. Las tendencias actuales incluyen la proliferación de nuevas pruebas, estándares más altos, tecnología mejorada, incremento de la objetividad, mayor conciencia pública, la computarización de las pruebas y su aplicación por Internet.

En cuanto al futuro, cualquier cosa es posible, en especial en un campo tan controvertido como el de las pruebas. En la actualidad, la psicología está equiparada mejor que nunca en lo que se refiere a técnica, metodología, datos empíricos y experiencia, y los miembros de este campo nuevo y en crecimiento, como grupo, son relativamente jóvenes. Por lo tanto, no parece que sea poco realista o demasiado optimista, esperar que en los próximos 50 años veamos avances similares a los producidos en los pasados 50. Por otro lado, la psicología ha llegado tan lejos en los últimos 50 años que un avance comparable en los próximos 50 podría fácilmente producir resultados inimaginables hoy en día. Lo que suceda con las pruebas en el futuro dependerá de las metas y objetivos elegidos por aquellos en el campo y por su persistencia y creatividad para alcanzar sus metas.

8.5. RESUMEN: EVALUACIÓN INFORMATIZADA, EVALUACIÓN DE PROGRAMAS Y PREDICIONES DE FUTURO

A modo de resumen de lo desarrollado en este capítulo donde se enfatizan dos líneas de futuro en evaluación psicológica, se procede, en primer lugar a sintetizar la situación actual de la evaluación psicológica informatizada así como sus limitaciones (Cuadro 49), y en segundo lugar, lo relativo a las bases conceptuales y metodológicas en el diseño y evaluación de programas, intervenciones o tratamientos (Cuadro 50). Para concluir se perfilan algunas predicciones de futuro sobre la evaluación psicológica y sobre el papel del psicólogo en España (Cuadro 51).

Cuadro 49. La evaluación psicológica informatizada

LA EVALUACIÓN PSICOLÓGICA INFORMATIZADA	
Situación actual	• En la actualidad los ordenadores se han extendido a múltiples campos y aspectos de la vida cotidiana, y este impacto se ha visto también reflejado en la Evaluación Psicológica. • El empleo del ordenador ha sido debatido ampliamente, manifestándose posturas contrapuestas entre los psicólogos. Pese a las críticas, estas herramientas, aunque en el futuro irán perfeccionándose, actualmente pueden ser una ayuda en el proceso que implica la Evaluación Psicológica. • La incorporación del ordenador como instrumentación auxiliar en Evaluación psicológica tiene lugar en relación a tres momentos del proceso evaluador: 1) En la administración, 2) En la corrección-puntuación, y 3) En la interpretación de los resultados.
Administración informatizada	Aplicación informatizada de instrumentos tradicionales Actualmente se pueden aplicar por ordenador muchos instrumentos tradicionales: • Historias Clínicas: Existen numerosos programas distintos para la realización de Historias Clínicas. Pueden ser cumplimentadas por el psicólogo, por los padres del sujeto, o por el sujeto mismo. • Cuestionarios e Inventarios Conductuales: Dirigidos a la evaluación de diversos problemas como ansiedad, estrés, depresión, dolor crónico, alcoholismo, abuso de fármacos, problemas sexuales... • Técnicas Subjetivas: Listas de adjetivos y Técnicas de Rejilla. • Tests de Personalidad e Intereses (MMPI, 16 PF, Strong, EPI....). • Baterías de Aptitudes (PMA), Inventarios de conducta adaptativa, Tests de rendimiento,

	Tests de inteligencia... Cada vez están apareciendo más procedimientos para informatizar instrumentos tradicionales: • Programas que permite informatizar cualquier test de papel y lápiz, tanto autoinforme como test de ejecución (Cuestionarios Individuales Informatizados de Calero y Araque, 1991a, 1992). • MACI (Calero y Araque, 1990a): Matriz para informatizar cuestionarios. • Visual Basic 6.0 para Windows. Cuestionarios de personalidad por ordenador (Pelechano y Darías, 2000)... Aplicación de instrumentos para ser administrados únicamente por ordenador DRIVERTEST (Monterde y Bort, 1988) homologado por la Dirección General de Tráfico: coordinación visomotora, velocidad de anticipación, tiempos de reacción, atención concentrada y resistencia vigilante a la monotonía. HYPERCARD (Fernández et al., 1992) para evaluar la inteligencia espacial Aplicaciones combinando ordenador y vídeo SHIPS (Loeding y Crittenden, 1994): Orientación vocacional de personas sordas. Mide habilidades interpersonales relacionándolas con la elección profesional. Tests adaptativos computerizados (TAI) • Un TAI es una estrategia informatizada de evaluación con la que se pretende estimar el nivel de habilidad (entendido en sentido amplio: como destreza, rasgo, aptitud, conocimiento...) de una persona de la manera más precisa y eficiente posible. Para realizar un TAI se necesita un banco de ítems calibrado, es decir, con propiedades psicométricas conocidas y obtenidas desde un modelo TRI y un algoritmo informático que controla la aplicación del test. Comienza la aplicación de un test a un sujeto por un ítem cualquiera (normalmente de dificultad media) y en función de las respuestas del sujeto el algoritmo va decidiendo el ítem siguiente a aplicar, hasta llegar al último, a partir del cual se estima el nivel de habilidad según la función de estimación asociada a ese ítem. • TAC-g es un instrumento de medida del factor "g" de inteligencia mediante ítems de las matrices de Raven (Palacios et al., 1994). • Morfo Bank: medición de una habilidad lingüística (Renom y Puyuelo, 1994). • TRASl: Test adaptativo informatizado para la evaluación del razonamiento secuencial y la inducción (Rubio Franco y Santacreus, 2003) • La escala de ajuste emocional "EAE". Parte de un cuestionario de personalidad desarrollado a medida para fines de selección de personal (CPSP) (Aguado et al., 2005). Instrumento auxiliar en la aplicación de técnicas de potencial de aprendizaje • Sistemas de evaluación-entrenamiento programas informáticos diseñados para suministrar al sujeto, de forma controlada, señales y reglas dirigidas a la solución de la tarea. • DynaMath (Gerber et al. 1994): sistema de evaluación dinámica de multiplicación de múltiples dígitos mediante ordenador. • EPA-TEST (Calero y Araque, 1990b) - EPA-2, Fernández-Ballesteros et al., 2000: evaluación del potencial de aprendizaje.

	Desarrollo de Programas Informáticos Integrales • <i>PPD, Psychological Testing Program</i> (Zaal y Pieters, 1987): Programa que incluye de 14 a 16 pruebas (tests perceptivos, verbales, cuestionarios de personalidad, valores interpersonales...), permitiendo diversos modos de presentación de la tarea, diversas tareas y diversos criterios normativos. • <i>Vienna System</i> (TEA, 1992): Para diagnóstico o terapia, incluye pruebas de: memoria visual, vigilancia, atención continuada, detección de señales, rendimiento en series de trabajo, matrices progresivas, aptitud espacial, coordinación bimanual... • <i>TEA System 4.0</i> (TEA, 2005): Permite aplicar diversos tests de personalidad (CAQ, CPI, MMPI-2, 16PF-5, HSPQ...), evaluar resultados, definir los parámetros de cada test y ejecutar variadas funciones. Aplicación de pruebas basadas en criterios (pruebas de competencia mínima) Denominadas pruebas de competencia mínima permiten determinar la posición del individuo con respecto a un dominio de conducta previamente definida.
Corrección-puntuación informatizada	1. Corrección mecánica de tests de plantillas, que actualmente se pueden realizar para gran cantidad de tests, mediante la utilización de hojas de respuesta especiales y lectoras ópticas. 2. Programas dirigidos a valorar y puntuar las respuestas de los sujetos a técnicas consideradas de corrección conflictiva por la falta de estructuración de sus criterios valorativos: test de Rorschach, tests proyectivos de dibujo, test giestáltico visomotor de Bender, batería de Luria... 3. Programas para construir y aplicar un banco de normas, que permita la comparación de las puntuaciones obtenidas por un sujeto con cualquier criterio normativo o población de referencia previamente establecido. • Datacrhono: para registrar datos de observación en situaciones naturales que permite el volcado posterior en un ordenador para su análisis e interpretación. • Sistema EBASS: para registrar variables de conducta de los estudiantes y variables ambientales de la situación de enseñanza (Greenwood y otros, 1994). • Registros psicofisiológicos: para el análisis de la presión sanguínea... (Reyes del Paso, 1991; Demster, 2001). Algunos programas informáticos en el mercado español • <i>Info-WAIS III</i> (Seisdedos, Fluiters y Carrasco, 2004): Programa informático en Windows para ayudar a la corrección e interpretación de la escala WAIS-III. • <i>RR</i>. Sistema experto para la interpretación del Test de Rorschach (Drag & MCF, 1995; García Polín, 1996): Es un sistema concebido como ayuda para la interpretación del test de Rorschach. Basado en un potente sistema experto, aprende con los sucesivos casos y se maneja fácilmente. • <i>RECORD</i>. La técnica de la Rejilla (Feixas y Cornejo, 1992): Es un instrumento informatizado que permite evaluar las dimensiones y estructura psicológica del constructo "significado personal", aspectos de la forma en que una persona da sentido a su experiencia y organiza su mundo. • <i>PPI</i> (Sebastian y Machlab, 1994): Programa informático para evaluación de la personalidad a partir del test 16 PF. Dispone de una base de datos interactiva y elabora un perfil de la personalidad y un completo informe sobre la misma y sus posibles desajustes, incluyendo coeficientes de correlación con otros casos.

	• Info-WISC-R (Seisdedos, 1993): Programa de ayuda a la corrección e interpretación de la escala WISC-R. Se introducen las puntuaciones directas obtenidas por el sujeto y el ordenador ofrece las puntuaciones transformadas de acuerdo con el manual, perfil del niño, y un informe para uso del psicólogo con los comentarios pertinentes. • SOCGRAMA (Almari y Gil, 1993): Consiste en un programa informático para el tratamiento de los sociogramas que, hasta ahora, resultaba muy laborioso a la hora de tratar y tabular los datos. • PSIQUE CP (PSICWARE, 1993): Es un sistema informático para gestión de un gabinete de Psicología Clínica con varios terapeutas. Permite realizar la gestión de los pacientes, conservar historias clínicas, y la aplicación de diversas pruebas de evaluación de uso frecuente en el campo clínico (EPQ, STAI, STAIC...), así como la obtención de informes sobre los resultados del sujeto. • METRIX Engine (Renom, 1996): Sistema de análisis y evaluación de tests. Facilita la gestión e interpretación de resultados en bases de datos y aporta las bases metodológicas y técnicas para la elaboración de tests de todo tipo. • CPTII. Test de atención de Conners. Evaluación computerizada de la atención • VISIONLAB: Sistema de gafas con obturador conectadas a un ordenador que permite la presentación de estímulos visuales en tres dimensiones. El ordenador puede funcionar como taquistoscopio, tambor de memoria analizador de percepción visual... • Test de golpeo digital computerizado: Test de golpeo utilizando el teclado del ordenador para medir la coordinación. • DATALAB2000: Dieciséis canales de evaluación de variables fisiológicas: ECG, EMG, EEG, GSR, presión sanguínea, conducción nerviosa, tensión muscular, ritmo cardíaco, respiración, fuerza, estimulación muscular... • Sistema Poligráfico Computerizado: Cuatro canales para la evaluación fisiológica: respiración, respuesta galvánica de la piel, presión sanguínea/ritmo cardíaco. • BIOPAC: Sistema modular de biofeedback y control fisiológico de doble canal. • BEI. Batería Escolar Informatizada (TEA, 2008): Batería de pruebas para edades de 6 a 18 años, que evalúa aptitudes intelectuales, personalidad, técnicas de estudio, intereses profesionales...
Interpretación de los resultados informatizada	Tipos de informes en la interpretación informatizada (Cohen y Swerdlik, 2001): • <i>Informes de calificación (simples y/o extensos)</i>: Suministran puntuación por escala y perfiles. Los extensos informes de calificación proporcionan criterios de significación estadística de puntuaciones o de relaciones entre puntuaciones. • <i>Informes de interpretación</i>: Descriptivos, tamizados y/o consultivos. Se acompañan de una narración que puede ser, en el primer caso, una descripción breve y concisa de cada una de las escalas. En el segundo incluye relaciones entre escalas, señalando aquellas que puedan ser interesantes para la interpretación. Los consultivos suministran un informe detallado en lenguaje adecuado para profesionales. • <i>Informes integradores</i>: Tratan de combinar datos de diferentes fuentes, algunos incluso de diferentes modelos teóricos para su interpretación. Dentro de este grupo se pueden incluir los <i>sistemas expertos</i>. <i>Sistemas expertos</i>: transferencia de los conocimientos y métodos de la inteligencia artificial a la resolución de problemas de diagnóstico y tratamiento. • CATCEC (Computer-Supported Assessment and Treatment Consultation for Emotional Crises) (Hedlund et al., 1987): Ayuda a la evaluación y tratamiento de las crisis emocionales. • DAI. Sistemas expertos para la evaluación del autismo (Adarraga y Zacagnini, 1992).

	• PSYMEDIA y DIASYS (Westmeyer y Hagebock (1992): Para formalizar el proceso evaluativo, permiten auxiliar el trabajo del evaluador a través de tal proceso. El PSYMEDIA es un sistema de inferencias que ayuda al evaluador a elaborar un diagnóstico con datos de pruebas basadas en la teoría clásica de los tests. El DIASYS 1 es un sistema aplicable a la evaluación psicoeducativa. • SMH: PAL (Hofmeister et al. 1994): para identificar procedimientos de tratamiento para estudiantes con dificultades. Ayuda al profesor a identificar intervenciones que pueden ser utilizadas con estudiantes que exhiben determinados problemas de conducta. • IPFSE (Intelligence Psychological Disorders Symptoms Evaluator): Sigue el formato de diagnóstico del DSM-IV (Cheng y Liou, 2003).
Ventajas de la evaluación informatizada	• Ahorra tiempo al profesional. • Disminuye costes. • Incrementa la fiabilidad por la mayor estandarización en la toma de datos. • No se aprecian diferencias en la cantidad de información recogida, ni en la deseabilidad social en los cuestionarios de personalidad y actitudes... • Aumenta la percepción de confidencialidad y conduce a menor cantidad de respuestas sesgadas. • Su carácter impersonal puede favorecer un mayor sinceramiento en sujetos muy sensibles a quienes les resulta más fácil admitir sus dificultades a una máquina que al entrevistador. • Las entrevistas administradas por ordenador recogen información clínicamente significativa de pacientes adultos en el contexto de servicios de salud mental. • Se acumulan datos de diferentes sujetos, se construyen baremos propios, se tienen en cuenta patrones de respuesta que resulten de interés.
Problemas de la evaluación informatizada	Aplicación informatizada • Equivalencia de formatos (tradicional-informatizado): Necesidad de contar con una mayor verificación empírica de ambas versiones del mismo test. • Despersonalización-Deshumanización de la tarea: Desvirtua la complejidad del proceso al trasmitirle al sujeto una imagen de la evaluación en lo que lo central es el instrumento y no el grado de compromiso y la responsabilidad del psicólogo en su relación personal con el entrevistado. • La falta de entrenamiento adecuado y falta de experiencia del evaluador. • Una actitud negativa o ansiosa frente a los ordenadores por parte de los evaluados. • El aumento de la accesibilidad a los tests por parte de clientes potenciales podría provocar un entrenamiento inadecuado. Interpretación informatizada • Atenta contra la centralidad del rol del psicólogo en la tarea de integración de los datos de la evaluación. • Aún no ha logrado su madurez científica: Validez de la narración generada por el ordenador como juicio clínico sobre un sujeto en particular no ha sido establecida, y además, este tipo de descripciones suelen ser tan generales que podrían ser aplicadas a infinidad de sujetos. • Peligro de su uso inadecuado por personas que aceptaran de forma poco crítica el contenido del informe narrativo. Actualmente existen gran cantidad de tests informatizados de todas las técnicas, sin embargo su uso permanece constante desde el inicio de los noventa. Algunas razones explicativas de este hecho pueden ser:

	1. Falta de entrenamiento y formación de los psicólogos en el uso de tests computerizados. 2. Alto coste económico que supone mantener los equipos y programas actualizados. 3. Dudas sobre la equivalencia de la administración tradicional e informatizada. 4. Resultados incongruentes de la interpretación. El ordenador, a veces, produce una incongruencia entre interpretaciones. La evaluación informatizada deja fuera otros elementos no verbales (gestos faciales, tono de voz...) y anulan la interacción evaluador-evaluado que puede tener cierta influencia en la ejecución. El uso del ordenador en administración y puntuación, aún con reservas, es, hoy, aceptado. 5. Problemas de validez: • Las interpretaciones informatizadas pueden ser excesivamente generales • En uso del sistema de expertos es una buena estrategia inicial para garantizar la validez de un programa pero insuficiente. • Aunque los informes focalizados en el uso de criterios externos son los realmente demostrativos de validez, hasta la fecha no han arrojado resultados contundentes. Las muestras de validación del formato informatizado suelen ser pequeñas, y las medidas externas de criterio suelen ser inadecuadas. Existen escasas investigaciones sobre la consistencia de los informes, e incluso algunas de éstas detectan discrepancias entre informes, lo que pone de relieve los problemas de validez. 6. Problemas éticos ya que los datos de los clientes en un ordenador no están protegidos adecuadamente. Conclusión: En la actualidad el informe computerizado debe ser un auxiliar

Cuadro 50. La evaluación de programas, intervenciones o tratamientos

LA EVALUACIÓN DE PROGRAMAS, INTERVENCIONES O TRATAMIENTOS	
Definición	• Sistemática investigación del mérito, valor o éxito de concretos tratamientos, intervenciones o programas llevados a cabo con el fin de producir cambios comportamentales. • Investigar los resultados de una intervención.
Criterios	• Útil • Realista y factible • Ética • Precisa
Modelos	La evaluación formativa: • Se realiza sobre el proceso de implementación de una determinada intervención, tratamiento o programa. • Tiene como objetivo mejorar el programa en curso de elaboración. • El evaluador debe ser preferentemente el propio implementador del programa, un evaluador-implementador altamente conocedor del programa en el aspecto cognitivo y emocional.

	• Implica evaluación continua: Se realiza a lo largo y en distintos y sucesivos momentos de la implementación del programa con la finalidad de ejercer la función de corrección del proceso. La evaluación sumativa: • Se refiere a los resultados derivados de la aplicación de un programa. • Tiene como objetivo determinar el valor de este programa cuando su implementación ha sido acabada. • El evaluador debe ser una persona externa e independiente del proceso que se va a evaluar. • Evaluación final: Se considera el producto al final, teniendo por finalidad la sanción social o determinación del grado de precisión o éxito global alcanzado.
Fases y Tareas	FASE 1. DISEÑO DEL PROGRAMA DE INTERVENCIÓN • Justificación y antecedentes del programa: Razones por las que se va a llevar a cabo, interés del programa, ventajas-utilidades del programa que va a diseñarse y antecedentes del programa. • Fundamentación teórico-empírica del programa: Características y necesidades en la etapa evolutiva para la que está diseñando el programa; definición de conceptos teóricos relevantes al programa; explicaciones teóricas propuestas por las distintas perspectivas epistemológicas de la psicología sobre los conceptos teóricos; factores que influyen en las variables de intervención; investigaciones previas similares o complementarias, así como procedimientos y resultados obtenidos en ellas. • Objetivos del programa: Identificar objetivos generales que se pretenden con el programa así como objetivos específicos, concretos incluidos en estos objetivos generales. • Características del programa: Determinar a quién va dirigido, donde se va a realizar, quién lo administrará, durante cuanto tiempo... • Configuración del programa: Identificar los módulos del programa y el número de actividades que contiene cada módulo o sección, las técnicas que subyacen a las actividades y al programa en su conjunto... • Actividades o acciones del programa: Las actividades constituyen la esencia de un programa por lo que se requiere describir detalladamente las actividades seleccionadas para cubrir los objetivos, las actividades que configuran el programa deben describirse de forma homogénea, por ejemplo, en formato de ficha técnica (objetivos de la actividad, descripción de la actividad e instrucciones para los participantes, preguntas para la discusión o debate posterior a la actividad, materiales necesarios para ejecutarla...). • Procedimiento para la implementación del programa: Redactar la consigna o encuadre de introducción al programa, elaborar un cuadro o esquema que determine la temporalización de la secuencia de las acciones a realizar para aplicar el programa, tiempo que durará la implementación, número de sesiones de intervención que se llevarán a cabo, tiempo de duración de cada sesión, actividades que se realizarán en cada sesión... • Factores de influencia en el programa: Factores o variables que pueden afectar positiva o negativamente a la implementación del programa, por ejemplo, la formación de los

	adultos que implementarán el programa, el espacio físico donde se realice... FASE 2. DISEÑO DE LA EVALUACIÓN DEL PROGRAMA • Finalidad de la evaluación y factibilidad: En primer lugar, el evaluador debe clarificar la finalidad de la evaluación del programa, si la evaluación que se propone es factible de llevarse a cabo y qué dificultades podría encontrarse durante la evaluación. • Identificar las variables dependientes y los instrumentos de evaluación con los que se medirán estas variables: En este punto el evaluador debe preguntarse, en primer lugar, qué variables dependientes se van a medir teniendo en cuenta la finalidad de la evaluación, cuales deben ser las variables objeto de análisis para evaluar el programa que permitan mejor evaluar los cambios, y si existen otros efectos del programa que pueden ser evaluables. En segundo lugar, debe identificar con qué instrumentos de evaluación medirá las variables dependientes seleccionadas como objeto de evaluación. Deberá valorar si existen instrumentos estándar con garantías psicométricas de fiabilidad y validez o se requiere la creación de instrumentos "ad hoc" para medir estas variables. Así mismo se preguntará por las fuentes que ofrecen mayores garantías para la recogida de datos, es decir, quién aportará la información de mayor calidad para la evaluación de los efectos del programa. Se deberá determinar el procedimiento a seguir para administrar la evaluación. Así mismo se deberán identificar factores o variables extrañas que pueden afectar la recogida de datos en la evaluación, por ejemplo, formación de los evaluadores, o factores derivados de los instrumentos o del contexto que pueden afectar a los datos que se recogen... • Selección del diseño de investigación a utilizar para evaluar el programa: La primera decisión será elegir entre los dos macromodelos de evaluación, el modelo no lineal (evaluación formativa), o el modelo lineal (sumativa) y para ello se deberá tener en cuenta la finalidad de la evaluación (evaluación del resultado o evaluación del proceso). No obstante, en ocasiones, y siempre teniendo en cuenta los objetivos de la evaluación la utilización de ambos modelos puede complementarse. Tras la elección primaria del macromodelo se elegirá el diseño de investigación a utilizar pudiendo optarse o bien por diseños experimentales (caso único, diseños clásicos de grupos pretest-postest con o sin grupo de control...), o bien por diseños no experimentales como la descripción naturalista mediante técnicas observacionales. Finalmente cabe plantearse si existen amenazas a la validez interna y externa, y en ese caso determinar procedimientos para el control de estos sesgos. • Calendario para la evaluación: Cuando todas estas cuestiones se han determinado se procede a preparar el calendario de la evaluación, es decir, la secuencia con la que se procederá a la recogida de los datos. FASE 3. IMPLEMENTACIÓN Y EVALUACIÓN DEL PROGRAMA • Implementación del programa y recogida de datos para la evaluación: Cuando el programa y la metodología de evaluación han sido diseñadas, se procede a la implementación del programa y a la recogida de datos para llevar a cabo su evaluación. • Análisis de los datos: El análisis de datos dependerá de la naturaleza de los datos recogidos y del tipo de preguntas que hayan sido formuladas en relación al programa. No será lo mismo si se trabaja con datos cuantitativos o cualitativos, así como si se requieren resultados descriptivos del programa (frecuencias, porcentajes, medias, desviaciones típicas) o si se buscan nexos causales que implican análisis inferenciales (análisis de varianza univariados o multivariados, regresión múltiple...). En esta fase el evaluador se

	pregunta sobre el soporte en el que va a almacenar la información obtenida, sobre el tipo de análisis que se pueden realizar con los datos que se han recogido y para responder a las preguntas a las que se pretende dar respuesta con la evaluación. El criterio principal que debe guiar el análisis a la hora de evaluar un programa es la identificación de los efectos del programa, por un lado, su eficacia (si se han cumplido los objetivos del programa), y por otro, su efectividad (si ha tenido otros efectos positivos independientes de los objetivos). No obstante, cabe destacar que el análisis de los datos no debe consistir en la aplicación mecánica y ciega de las técnicas previstas en el diseño, sino un proceso mucho más reflexivo e interactivo en el que cada resultado puede influir en las decisiones acerca de los siguientes análisis a llevar a cabo. FASE 4. ELABORACIÓN DEL INFORME DE LA EVALUACIÓN DEL PROGRAMA • <i>Portada del Informe:</i> Debe contener la denominación del programa evaluado, el nombre del evaluador o evaluadora que lo ha realizado, así como la fecha de redacción del informe. • <i>Resumen:</i> Debe incluir lo esencial de la evaluación, es decir, los objetivos de la evaluación, una breve descripción del programa y de su procedimiento de implementación, una breve descripción de la metodología y los instrumentos de evaluación utilizados, así como los resultados más relevantes obtenidos. • <i>Objetivos e hipótesis del estudio de evaluación del programa:</i> Se deben explicitar los objetivos de la evaluación del programa, es decir, de la investigación llevada a cabo con él, y así mismo se deben describir todas las hipótesis que se han formulado sobre los efectos del programa, o sobre otras cuestiones asociadas a la evaluación... • <i>Diseño y procedimiento de la investigación:</i> Se informa del diseño utilizado en la investigación llevada a cabo para evaluar los efectos del programa, así como del procedimiento seguido, es decir, las fases y tareas realizadas en el transcurso del estudio. • <i>Participantes:</i> Se describe de forma detallada las características de la muestra con la que se ha llevado a cabo la evaluación del programa (número de participantes y características de los mismos) y el procedimiento de selección muestral utilizado. • <i>Descripción del programa de intervención:</i> Se informa sobre las razones por las que se lleva a cabo el programa e interés del mismo, la fundamentación teórico-empírica sobre la que se sustenta (conceptos, teorías, estudios...), los objetivos generales y concretos del programa, las características, la configuración, el procedimiento de la implementación llevado a cabo describiendo las actividades o acciones realizadas... • <i>Metodología e instrumentos de evaluación del programa:</i> Se aporta información sobre la metodología de evaluación empleada, se describen los instrumentos de evaluación utilizados (qué variables evalúan, como se han aplicado, corregido e interpretado los datos que se registran con éstos, garantías científicas de los instrumentos: datos de fiabilidad y validez...), la secuenciación llevada a cabo en la evaluación (sesiones de evaluación y temporalización de las mismas), variables extrañas controladas para garantizar la validez de los datos recogidos... • <i>Resultados:</i> Se explican los análisis realizados con los datos para responder a las preguntas planteadas y se exponen los resultados obtenidos de forma ordenada. Se presentan los resultados cuantitativos en tablas y gráficos, así como los resultados cualitativos complementarios.

	• Conclusiones: Se clarifica en qué medida se han encontrado resultados significativos que confirman la eficacia (ha cumplido los objetivos que el programa se proponía) y/o la efectividad del programa (ha tenido otros efectos positivos independientes de los objetivos). Por un lado, se sintetizan los resultados más relevantes y se aportan explicaciones a los mismos (factores o variables que han desempeñado un papel determinante en la obtención de estos resultados...), se contrastan los resultados con las hipótesis propuestas evidenciando si éstas se confirman o refutan, y también se contrastan los resultados con los obtenidos en otros estudios que o bien se han realizado con el mismo programa o bien se han llevado a cabo con programas que tenían objetivos similares. Por consiguiente, se procede a la presentación de las principales conclusiones del trabajo las cuales se debaten a la luz de parámetros teóricos y empíricos. • Recomendaciones: Teniendo en cuenta los hallazgos encontrados en el estudio de evaluación del programa, para finalizar el informe se procede a formular un conjunto de recomendaciones que pueden permitir mejorar el programa y por consiguiente sus efectos. Por un lado, se pueden identificar diversos aspectos positivos que deberían permanecer en subsiguientes implementaciones, tanto en lo referente al programa como a su evaluación, y por otro, se deben recomendar cambios que permitirían mejorar puntos débiles o bien del programa o bien de la metodología y los instrumentos de evaluación utilizados para evaluar sus efectos.
Diseños de Investigación	Los diseños para la evaluación de programas pueden variar desde el más riguroso experimento de campo a la descripción naturalista, sin olvidar, claro está, la diversidad de diseños cuasi-experimentales o los correlacionales. 1. Diseños experimentales y cuasiexperimentales: 1.1. Diseños clásicos de grupos • Diseño pretest-posttest con grupo de control equivalente • Diseño posttest con grupo de control equivalente • Diseño pretest-posttest con grupo de control no equivalente • Diseño de series temporales con grupo experimental sólo • Diseño de series temporales con un grupo de control no equivalente • Diseño pre-experimental antes-después 1.2. Diseños experimentales de caso único 2. Diseños No experimentales: • Diseños diacrónicos • Diseños Sincrónicos • Diseños mixtos Diacrónicos/Sincrónicos
Instrumentos de Evaluación	Las características que deben tener los instrumentos de evaluación de intervenciones: • Objetividad • Estabilidad • Exactitud, Precisión o Validez • Sensibilidad al cambio

	Instrumentos de evaluación cuantitativos • Entrevista estructurada. • Autoinformes: el sujeto se valora a sí mismo • Cuestionarios: sobre el nivel de satisfacción del programa, o sobre la percepción subjetiva del cambio por efecto del programa... • Indicadores estadísticos y sociales: número de toxicómanos, tasas de empleo... • Observación estructurada y/o sistemática: frecuencia de conductas observables... • Archivos: calificaciones, hojas de servicio... • Medidas de rastro y erosión: número de cigarros, botellas... • Escalas de apreciación: jueces (padres o profesores) califican al sujeto en una escala de estimación en relación a alguna conducta... • Pruebas de ejecución y de inteligencia: piezas insertadas, número de aciertos, tiempo de ejecución... • Pruebas de rendimiento: exámenes... • Cuestionarios: medidas de autoinforme sobre habilidades, nivel de satisfacción, o logros obtenidos con el programa... • Entrevistas estructuradas. Instrumentos de evaluación cualitativos • Entrevista abierta. • Observación no estructurada: diarios.... • Técnicas sociométricas. • Grupos de discusión. • Cuestionarios o encuestas abiertas. • Técnicas etnográficas: entrevista callejera... • Autobiografía • Análisis históricos: a través de distintos tipos de datos como periódicos o informes con los que se pretende establecer la historia pasada de un evento. • Método Delphi: expertos a los que se envían los cuestionarios y los resultados de los mismos...

Cuadro 51. Algunas predicciones futuras sobre la evaluación psicológica y el rol del psicólogo en España

• Incremento de la cooperación interdisciplinar con otros profesionales y especialización del psicólogo en campos concretos. • Mayor incorporación de los psicólogos en los servicios de salud. • Mayor conexión entre evaluación y tratamiento. • Incremento de tecnología evaluativa que disponga de garantías científicas. • Mayor sensibilidad del psicólogo hacia el código deontológico tanto en lo que se refiere a la evaluación psicológica como al resto de las actuaciones profesionales. • Desarrollo de una mayor formación multicultural, así como de instrumentos que eviten significativamente sesgos relacionados con la cultura...