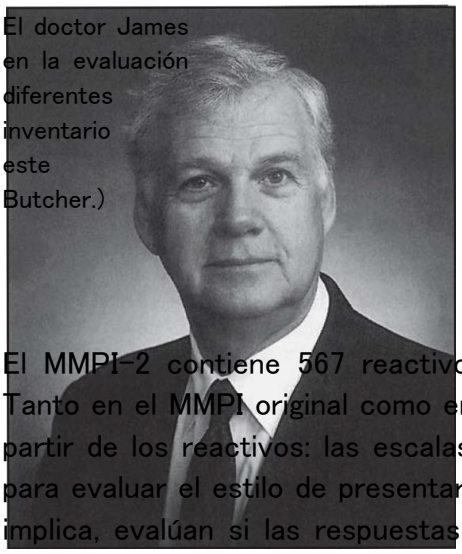
RE-ESTANDARIZACIÓN DEL MMPI: EL MMPL-2

A pesar de su extraordinaria popularidad, el MMPI no carecía de problemas. El MMPI original recibió varias críticas como: a) su dependencia de un sistema anticuado de diagnóstico psiquiátrico que había sido reemplazado por varias revisiones del DSM;
b) una muestra normativa que no era representativa de la población más amplia en términos de la diversidad socioeconómica; y e) reactivos que se habían vuelto anticuados o que eran objetables u ofensivos en su contenido. Esos problemas condujeron a un esfuerzo a gran escala, que inició en la década de 1980, para revisar el MMPI. Un comité de re-estandarización estableció varias metas importantes, entre ellas:

- Revisar y modernizar los reactivos del MMPI, incluyendo la eliminación de reactivos obsoletos u objetables (se cambió alrededor de 14% de los reactivos originales).
- Asegurar continuidad con la validez del MMPI original así como con las escalas estándar y complementarias.
- Desarrollar nuevas escalas para abordar problemas clínicos no cubiertos por el MMPI original.
- Obtener nuevas muestras seleccionadas al azar de adultos y adolescentes representativas de la población estadounidense y desarrollar normas apropiadas a la edad.
- Desarrollar nuevas distribuciones normativas para las escalas de adultos y adolescentes que reflejaran mejor los problemas clínicos.
- Obtener nuevos datos clínicos para usarlos en la evaluación y validación de las nuevas escalas.

La muestra normativa del MMPI-2 estuvo conformada por 2 600 sujetos, de 18 a 85 años, muestreados al azar en siete regiones de Estados Unidos. Se buscó equilibrar la muestra normativa en lo que se refiere a características demográficas y de género, tales como la pertenencia a un grupo étnico, para que fuera representativa de la población estadounidense. El MMPI-A (versión para adolescentes) se normó en más de 1 600 estudiantes de 14 a 18 años a lo largo de Estados Unidos. Las puntuaciones en el MMPI-2 y el MMPI-A se evaluaron mediante una comparación con la distribución de puntuaciones de esas nuevas muestras; es decir, se consideraba que una puntuación individual en el MMPI-2 era anormal si se desviaba de manera significativa de la media para esta muestra normativa. También se obtuvieron nuevos datos de una variedad de muestras clínicas.



El doctor James Butcher es profesor en la University of Minnesota. Él es una autoridad en la evaluación psicológica de la personalidad y de las conductas anormales en las diferentes culturas. Ha desempeñado un papel de liderazgo en el desarrollo del inventario multifásico de personalidad Minnesota-2 (MMPI-2) y en una versión de este instrumento para adolescentes (MMPI-A). (Fotografía cortesía de .lames Butcher.)

El MMPI-2 contiene 567 reactivos, aunque puede calificarse usando sólo los primeros 370. Tanto en el MMPI original como en el MMPI-2 se generan dos categorías amplias de escalas a partir de los reactivos: las escalas de validez y las clínicas. Las escalas de validez se utilizan para evaluar el estilo de presentar la prueba y la actitud del participante, y como su nombre lo implica, evalúan si las respuestas del examinado a los reactivos son válidas. Como advierten Butcher y Williams (1992), el MMPI es una medida de autorreporte y si los individuos desean distorsionar sus respuestas, pueden lograrlo con facilidad. Una presentación distorsionada como resultado de una forma aleatoria de responder o de un esfuerzo motivado por mirarse bien o

mal en el MMPI es una amenaza a la confiabilidad y la validez de la prueba. Las escalas de validez se diseñaron con el fin de evaluar la apertura, honestidad, cooperación y disposición del individuo a compartir información personal; las escalas clínicas del MMPI se elaboraron para evaluar el estado psiquiátrico y las características de personalidad del examinado.

Escalas de validez Las escalas de validez del MMPI-2 fueron desarrolladas a partir de la suposición de que ciertos patrones de respuesta a los reactivos de verdadero o falso representan respuestas no confiables o sesgadas que podrían invalidar las puntuaciones en las escalas clínicas. Por ello, en el MMPI-2 se incluyeron reactivos que en la población normal tienen tasas de ocurrencia extremadamente altas o extremadamente bajas (es decir, tasas base alta o baja). Es probable que los individuos que refrendan un número desproporcionadamente alto de esos reactivos estén tratando de dar una imagen muy favorable o muy negativa de sí mismos. En el MMPI-2 existen cuatro escalas de validez principales. La escala No poder contestar (?) es sólo el número total de reactivos dejados en blanco o calificados como verdadero y falso por el examinado. Un alto número de reactivos en blanco invalidarán al MMPI. La escala de mentiras (L) es una medida de la tendencia a presentarse de manera abiertamente favorable, es decir, afirmar que se es mucho más virtuoso, ético o moral que lo que marcan las normas de la población. Si bien un solo reactivo puede en realidad caracterizar a un individuo, la probabilidad de que un gran número de esos reactivos lo haga es muy baja. La escala de infrecuencia (F) es una medida de la tendencia a reivindicar un número exorbitante de problemas psicológicos o a exagerar los problemas de ajuste. La escala F se compone de reactivos que rara vez son corroborados por la muestra normativa o por la muestra de pacientes. La escala de defensividad (K) es un indicador de la defensividad en la prueba y se usa para detectar clientes que se describen en términos abiertamente positivos. Como una aproximación defensiva a la presentación de la prueba suprimirá, con toda probabilidad, las puntuaciones en las escalas clínicas, se agrega una porción de K a algunas escalas clínicas del MMPI-2. Finalmente, se han desarrollado varias escalas o puntuaciones nuevas de validez para usarse con el MMPI-2, incluyendo la tendencia a responder en forma inconsistente a reactivos o a dar respuestas indiscriminadas de verdadero o falso (Butcher et al., 1989).

ESCALAS CLÍNICAS Después de determinar que se ha obtenido un perfil válido, la puntuación total del participante en cada una de las 10 escalas clínicas del MMPI-2 se convierte a un puntaje T (puntuaciones con una media de 50 y una desviación estándar de 10) con base en la muestra normativa. Se considera que los puntajes T superiores a 65 (el percentil 92° de la muestra normativa) son clínicamente significativas, es decir, la significancia clínica se define en términos de desviación de la norma. Las 10 escalas clínicas se describen en la tabla 9.1. Cada una de las escalas clínicas parece ser una escala auto contenida, es decir, cada una podría ser interpretada por sí misma en referencia a la dimensión original de psicopatología diseñada a representar. Así, por ejemplo, una alta puntuación en la escala **Se** podría sugerir un diagnóstico de esquizofrenia, y una puntuación alta en la escala **D** podría sugerir un diagnóstico de depresión. No obstante, décadas de investigación con el MMPI y con su revisión han indicado con claridad que la interpretación de las puntuaciones de escalas únicas no es la forma más efectiva de usar esta medida. En lugar de eso, los examinadores consideran el perfil de puntuaciones en las 10 escalas clínicas al interpretar el MMPI-2.

INTERPRETACIÓN DE LOS TIPOS DE PERFIL DEL MMPI-2

La aproximación del análisis de perfil a la interpretación de las respuestas al MMPI-2 maximiza la cantidad de información utilizada. En lugar de considerar una escala a la vez, un psicólogo clínico puede interpretar el conjunto completo de respuestas en las escalas. Los tipos de

perfiles se definen por la puntuación más alta o por puntuaciones por encima de 65 T en las escalas clínicas. Así, un examinado que haya obtenido puntajes T de 82 en la escala D (escala 2) y de 76 en la escala Pt (escala 7) tendría un perfil 2-7. Se utilizan dos reglas generales para interpretar los tipos de perfil: el perfil es claramente definido por dos o más escalas que alcanzan una significancia interpretativa (es decir, por encima de una puntuación de 65 T), y ha existido investigación suficiente sobre las descripciones conductuales para el perfil. La interpretación de perfiles basada en las dos o tres escalas más altas es la aproximación de mayor uso (Graham, 2000; véase tabla 9.2). Los clínicos pueden buscar el perfil (por ejemplo, 2-7, 3-1-2) en cualquiera de una serie de libros de codificación MMPI-2 que describen la personalidad y la psicopatología características de los pacientes que obtuvieron un perfil particular.

INTERPRETACIÓN DE LAS ESCALAS DE CONTENIDO DEL MMPI

Se ha elaborado un segundo método para la calificación e interpretación del MMPI-2 basado en un enfoque que difiere del árido empirismo original que llevó al desarrollo de las escalas clínicas estándar. Los reactivos del MMPI-2 se agruparon en escalas de acuerdo con la validez de facie de contenido, y se desarrolló un segundo conjunto de escalas que reflejan diversos tipos de psicopatología. Las escalas de contenido incluyen reactivos que reflejan directamente los síntomas principales de los problemas clínicos (por ejemplo, ansiedad, depresión) así como varios aspectos de la vida diaria (por ejemplo, relaciones familiares). Por ejemplo, BenPorath, Butcher y Graham (1991) examinaron la habilidad de las escalas de contenido del MMPI-2 para distinguir entre pacientes internos diagnosticados con esquizofrenia y los diagnosticados con depresión mayor. Sus resultados indicaron que tanto las escalas clínicas originales como las nuevas escalas de contenido del MMPI-2 contienen información útil en el diagnóstico diferencial de esos dos trastornos. Más importante aún, las escalas de contenido poseen validez creciente con respecto a esta cuestión diagnóstica. Es decir, arrojan información mucho más útil que la obtenida de la calificación del perfil de las escalas clínicas.

TABLA 9.1
Escalas clínicas del MMPI-2

Escala 1: Hipocondriasis (H s). Esta escala se diseñó como una medida de las preocupaciones anormales o "neuróticas" sobre la salud corporal. El grupo criterio original lo componían pacientes que presentaban preocupaciones acerca de dolencias físicas para las cuales no había base orgánica. Los descriptores de personalidad asociados con una alta calificación en esta escala son: exhibir preocupaciones corporales excesivas, numerosos síntomas somáticos vagos y quejas indefinidas como malestar gástrico, fatiga, dolor y debilidad física; egoísta; centrado en sí mismo, pesimista; derrotista; cínico.

Escala 2: Depresión (D). Esta escala mide depresión sintomática, esto es, un marco mental generalmente negativo en el cual el individuo ha reportado baja moral, falta de esperanza en el futuro, insatisfacción con la vida y un bajo estado de ánimo. Los descriptores de personalidad asociados con puntuaciones elevadas son: deprimido, desdichado, disfórico, pesimista, autocrítico, proclive a la culpa, aletargado, carente de autoconfianza, introvertido, tímido, alejado de actividades, exhibe dificultad para tomar decisiones.

Escala 3: Histeria (Hy). Esta escala se desarrolló originalmente como una medida de lo que se conocía como histeria de conversión, actualmente denominada trastorno de conversión. Los individuos que califican alto en esta escala a menudo manifiestan un patrón inusual de características de personalidad compuestos por negación y exuberante asertividad social, pero cuando hay estrés en las relaciones el individuo puede verse incapacitado súbitamente por problemas físicos, muchas veces de origen desconocido. Los descriptores de personalidad son: reacciona al estrés desarrollando síntomas físicos como jaquecas, dolores de pecho, debilidad y taquicardia; desarrolla síntomas que aparecen y disminuyen con rapidez; carece de insight acerca de la etiología subyacente de los síntomas o motivos; no reporta agitación emocional grave; está centrado en sí mismo; espera mucha atención y afecto de los demás; se compromete en relaciones interpersonales superficiales e inmaduras.

Escala 4: Desviación psicopática (Pd). Esta escala es una medición de la conducta antisocial o conducta psicopática. Descriptores de la personalidad son: actividades antisociales, actitudes rebeldes hacia las figuras de autoridad, relaciones familiares tormentosas, tendencia a culpar a los padres de los propios problemas, una historia de bajo rendimiento en la escuela o el trabajo, impulsividad y un esfuerzo por obtener gratificación inmediata de los impulsos, un interés por los demás sólo en términos de cómo puede usárseles para satisfacer los propios propósitos.

Los individuos que califican alto en esta escala actúan sin considerar las consecuencias de sus acciones, tienden a no beneficiarse de la experiencia y se involucran en relaciones poco profundas y superficiales, sin embargo, son agradables y al principio causan una buena impresión.

Escala 5: Masculinidad-femineidad (Mf). Esta escala es diferente de las otras escalas estándar de varias maneras. El constructo que subyace a su desarrollo no es un síndrome clínico. Se diseñó originalmente para identificar rasgos de personalidad de "inversión sexual masculina" u hombres homosexuales que tenían un patrón de interés femenino. La escala Mf no es una escala de síntomas como la mayoría de las otras escalas estándar, ya que las elevaciones en esta escala

reflejan intereses, valores y características de personalidad. Los hombres que califican alto tienen problemas y preocupaciones sexuales; presentan conflictos con su identidad sexual; carecen de intereses masculinos estereotipados; tienen intereses estéticos y artísticos; tienen una orientación andrógena; son inteligentes, de pensamiento claro, lógicos; muestran buen juicio; son curiosos, creativos, imaginativos, sociables, sensibles a los demás. Las mujeres que califican alto rechazan los roles femeninos tradicionales; tienen intereses que son más estereotípicamente masculinos que femeninos; son activas, asertivas, competitivas, sociables, desinhibidas y tienen confianza en sí mismas.

Escala 6: Paranoia (Pa). Esta escala evalúa los patrones de conducta de suspicacia, desconfianza, conducta delirante, excesiva sensibilidad interpersonal, pensamiento rígido y exteriorización de la culpa comúnmente encontrado en los trastornos paranoides. Los descriptores de personalidad asociados con esta escala cambian considerablemente con una elevación incrementada, un patrón diferente de otras escalas. Los individuos con las puntuaciones más altas son francamente psicóticos, se involucran en un pensamiento trastornado, son delirantes, utilizan la proyección como mecanismo de defensa; mientras que los individuos con puntuaciones moderadamente altas manifiestan una predisposición paranoide, son demasiado sensibles y receptivos a las opiniones de los demás, racionalizan y tienden a culpar a los demás por sus problemas , son suspicaces y cautelosos, además de moralistas y rígidos en sus opiniones y actitudes.

Escala 7: Psicastenia (Pt). Esta escala se desarrolló originalmente para evaluar un trastorno psicológico que en la actualidad se describe como trastorno de ansiedad con rasgos obsesivo-compulsivos. Descriptores de la personalidad: los que califican alto son ansiosos, tensos , agitados, experimentan gran incomodidad, preocupación y sentimientos de aprensión; son muy nerviosos, deprimidos ; tienen dificultades para concentrarse; experimentan pensamiento obsesivo; se involucran en conductas compulsivas o ritualistas y en rumiaciones ; se sienten inseguros e inferiores; son autocríticos , autoconscientes ; establecen altos estándares de desempeño para sí mismos y para los demás; son perfeccionistas y concienzudos .

Escala 8: Esquizofrenia (Se). Esta escala reflejaba un intento por distinguir cuatro subtipos de esquizofrenia; el esfuerzo no tuvo éxito y los reactivos fueron agrupados en una sola escala. Descriptores de personalidad: quienes califican alto tienen un estilo de vida esquizoide; no se sienten parte del ambiente social; están aislados, alienados y malentendidos ; son ansiosos, rencorosos, hostiles y agresivos; tienen dificultades para separar la fantasía y la realidad; tienen dudas de sí mismos; pueden manifestar preocupación sexual; reaccionan al estrés refugiándose en la fantasía y en las ensoñaciones; están confundidos , desorganizados y desorientados; pueden reportar pensamientos o actitudes inusuales o alucinaciones; son aislados, recluidos, reservados ; evitan tratar con la gente y con las nuevas situaciones.

Escala 9: Hipomanía (Ma). Esta escala mide la conducta maniaca o hipomaniaca o la tendencia a actuar de manera eufórica, agresiva e hiperactiva. Descriptores de la personalidad: quienes califican muy alto son considerados como hiperactivos; hablan de manera acelerada; pueden padecer alucinaciones o delirios de grandeza; se involucran en actividad excesiva paro sin un propósito definido; efectúan

autovaloraciones poco realistas ; pueden tener demasiados proyectos en marcha a la vez; no utilizan la energía con prudencia; a menudo no concluyen los proyectos; tienen poco interés en la rutina y el detalle; tienen dificultad para inhibir la expresión de los impulsos.

Escala O: Introversión social (Si) Esta escala mide la incomodidad social, la inferioridad, la baja afiliación, la sensibilidad interpersonal, la falta de confianza y las quejas físicas. Descriptores de la personalidad: quienes califican alto son socialmente introvertidos; se sienten más cómodos solos; son reservados, tímidos, retraídos; son inseguros y se sienten incómodos en situaciones sociales; carecen de confianza en sí mismos; son sensibles a lo que otros piensen de ellos; son sumisos y dóciles en las relaciones interpersonales.

Nota; Adaptado de Butcher y William s (1992) y de Graham (2000).

CONFIABILIDAD Y VALIDEZ

La confiabilidad y validez de las escalas clínicas del MMPI se han examinado en cientos, si no es que en miles, de estudios. Tanto la consistencia interna como la confiabilidad test-retest de las escalas clínicas son razonablemente altas (por ejemplo, Parker, Hanson y Hunsley, 1988), en particular dados los cambios esperados en los síntomas de los pacientes con tratamiento. Se han expresado algunas preocupaciones de que muchas de las escalas clínicas correlacionan demasiado alto entre sí debido a los reactivos compartidos (por ejemplo, Butcher et al., 1989). Sin embargo, es importante darse cuenta de que diferentes formas de psicopatología no se distinguen perfectamente entre sí. En particular, la depresión y la ansiedad tienen tasas muy altas de co-ocurrencia o comorbiedad. En consecuencia, es razonable esperar también un traslape de los reactivos entre las 10 escalas clínicas del MMPI-2. La validez de las escalas del MMPI se ha probado sobre todo en estudios que examinan la habilidad de la prueba para discriminar entre grupos que se saben difieren en algún criterio conocido (es decir, validez discriminante). También en este caso la evidencia es sólida. Más aún, se ha demostrado que el MMPI-2 tiene validez creciente superior a la información obtenida con otros instrumentos de evaluación.

CRÍTICAS AL MMPI-2

El MMPI tiene una historia notable dentro de la psicología clínica, y el MMPI2 representa un progreso significativo sobre su predecesor. No obstante, a pesar de las fortalezas de este instrumento, ha recibido varias críticas. Algunas de las preocupaciones con el MMPI-2 son el resultado de la decisión de retener muchos rasgos de la versión original en aras de mantener la continuidad; otras involucran problemas que han surgido en el desarrollo de la nueva versión. Por ejemplo, no se hizo esfuerzo por adaptar el MMPI-2 a la teoría actual en el área de la personalidad y la psicopatología. Como resultado del uso de las escalas clínicas originales, las puntuaciones principales que se obtienen reflejan una conceptualización anticuada de la psicopatología. Varios de los nombres de las escalas (por ejemplo, histeria, psicastenia) se refieren a conceptos anticuados en psicopatología y pueden carecer de relevancia para las decisiones diagnósticas actuales. Además, a pesar del esfuerzo dedicado a obtener datos de una muestra normativa más representativa para el MMPI2, subsisten preocupaciones acerca del uso de las nuevas normas para generar y evaluar los perfiles o tipos de código específicos (Helmes y Reddon). Es decir, buena parte de la investigación sobre las características de los individuos con tipos particulares de perfil fue conducida usando el MMPI y sus normas originales. Aún no está claro si los mismos perfiles en el MMPI-2 necesariamente asegurarán las mismas interpretaciones que fueron apropiadas en la versión original.

TABLA 9.2 Ejemplos de interpretaciones de perfiles del MMPI-2
Butcher y Williams (1992) reportan el uso de códigos tipo para obtener información acerca de síntomas y conductas, características de personalidad y predicciones o disposiciones. Por ejemplo, un código tipo 2-7 (depresión y psicastenia) "parecerían ansiosos, tensos, nerviosos y deprimidos. Reportarían sentirse desdichados y tristes, y tenderían a preocuparse excesivamente. Se sienten vulnerables a amenazas reales e imaginarias, y suelen anticipar problemas antes de que ocurran, a menudo reaccionan en exceso a las tensiones menores como si fueran una gran catástrofe... parecen dóciles y pasivos-dependientes en las relaciones... usualmente muestran capacidad para crear vínculos emocionales profundos y tienden a apoyarse en la gente en un grado excesivo... suele diagnosticárseles como depresivos, obsesivo-compulsivos o con trastornos de ansiedad." El código tipo 2-8 (depresión y esquizofrenia) "parecen ansiosos, agitados, tensos y nerviosos. A menudo reportan tener problemas de sueño y no pueden concentrarse. Afecto perturbado y síntomas somáticos usualmente caracterizan la imagen clínica. Con frecuencia están clínicamente deprimidos, tienen un habla y pensamiento suaves y lentos... Suelen advertirse problemas con la ira y las relaciones interpersonales ... son abiertamente sensibles a las reacciones de los demás, suspicaces de las motivaciones de los otros, y pueden tener una historia de haber sufrido daño emocional... dependientes, poco asertivos, irritables y rencorosos ... A menudo temen perder el control de sus emociones ... se dejan llevar por la culpa y son auto-punitivos ... los diagnósticos más comunes son psicosis maniaco-depresiva , esquizofrenia, tipo esquizo-afectiva o trastorno de personalidad grave... a menudo se preocupan con pensamientos suicidas, y pueden tener planes específicos para quitarse la vida."

Nota: Tomado de Butcher y Williams (1992).

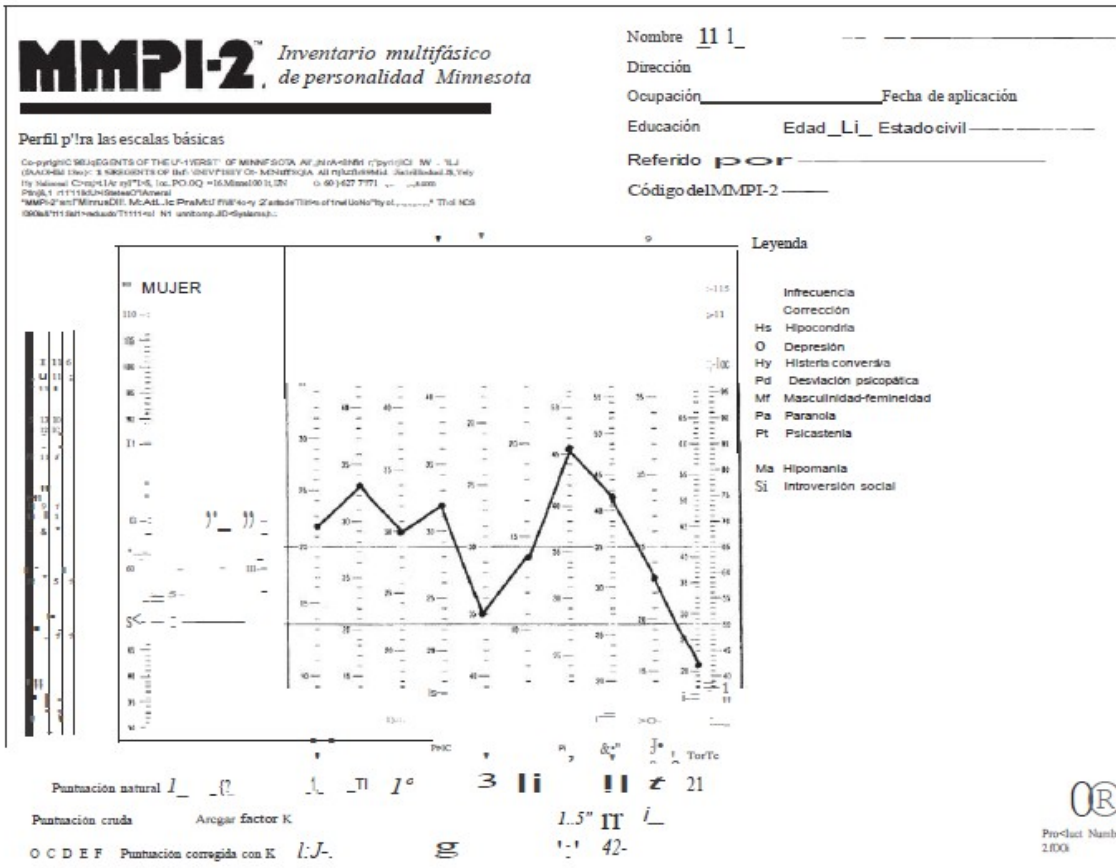


Figura 9.1
Perfil del MMPI-2 que representa las respuestas de Allison

EL INVENTARIO DE PERSONALIDAD NEO REVISADO

Una escala objetiva de personalidad desarrollada en fechas más recientes es el inventario de personalidad NEO revisado o NEO-PI-R (Revised NEO Personality Inventory), una prueba que tiene importantes diferencias con el MMPI. Costa y McCrae (1992) elaboraron el NEO-PI-R para evaluar los tres primeros de los "Cinco grandes" factores de personalidad (neuroticismo, extraversión y apertura, de ahí que en inglés se eligiera el nombre NEO) y posteriormente lo extendieron para incluir los cinco factores (N, E, O más conformidad y escrupulosidad). El NEO-PI-R es una medida de 181 reactivos con formas de autorreporte y para el observador. Los reactivos se responden en una escala de cinco puntos, que van de fuertemente en desacuerdo a fuertemente de acuerdo.

El formato y contenido de los reactivos del NEOPI-R son muy diferentes de los del MMPI-2. Los reactivos del NEO-PI-R tienen validez de facie o son obvios en su contenido; es decir, piden a los examinados que informen en forma directa acerca de las características que pretenden medir. Costa y McCrae (1992) sugieren que existen tres formas principales en las cuales pueden interpretarse las puntuaciones de la escala de personalidad resultante. La más directa es verlos como mediciones o estimaciones de cómo es en realidad el individuo. Una segunda interpretación más sutil es que representan el autoconcepto del individuo, o la forma en que se ve a sí mismo. Tercero, las puntuaciones pueden verse como auto presentación, o la forma en que el individuo desea ser visto por los demás, sea en general o en la ocasión específica en que se aplicó la prueba. Es probable que las tres interpretaciones sean correctas en cierto grado.

Costa y McCrae (1992) informan que el NEOPI-R tiene razonable consistencia interna y

confiabilidad test-retest, también informan que las correlaciones entre las auto calificaciones en el NEO-PI-R y las calificaciones de los cónyuges van de .43 (escrupulosidad) a .60 (apertura). Trull, Useda, Costa y McCrae (1995) aplicaron el NEO-PI-R a una muestra clínica y encontraron una estabilidad sustancial en las puntuaciones de personalidad a lo largo de un periodo de seis meses. También existe ahora un pequeño número de estudios que vinculan formas particulares de psicopatología con elevaciones en una o más de las escalas del NEO-PI-R. No obstante, esta medición no ha escapado de las críticas en el contexto de su utilización como herramienta de evaluación clínica (Butcher y Rouse, 1996). Por ejemplo, Ben-Porath y Waller (1992) plantearon dos preocupaciones acerca del NEO-PI-R. Primero, advirtieron que es poco probable que cinco factores amplios proporcionen información lo suficientemente específica para cumplir las necesidades de las aplicaciones clínicas. Y, segundo, criticaron la ausencia en el NEO-PI-R de escalas de validez como las del MMPI-2. Costa y McCrae argumentaron que los autorreportes de los individuos son, en general, fidedignos. Sin embargo, dada la naturaleza de las muestras y las situaciones en las que se emplea el NEO-PI-R para hacer evaluación clínica, la falta de escalas de validez es claramente una preocupación importante.

OTRAS PRUEBAS OBJETIVAS DE PERSONALIDAD

El MMPI-2 y el NEO-PI-R son dos de muchas pruebas objetivas de personalidad que se emplean en la actualidad en la investigación y la práctica clínica. Otras pruebas de gran uso incluyen el cuestionario multidimensional de la personalidad (MPQ, Multidimensional Personality Questionnaire; Tellegen y Waller, en prensa), el Inventario multiaxial clínico de personalidad de Millon (MCMI Clinical Multiaxial Personality Inventory; Millon, 1992, 1996), el inventario básico de personalidad (BPI, Basic Personality Inventory; Jackson, 1989) y el inventario de personalidad de California (CPI California Personality Inventory; Gough, 1990). Aunque todas esas pruebas tienen propiedades psicométricas adecuadas, el volumen de estudios empíricos que examinan la validez de esas mediciones es muy pequeño en comparación con el número de investigaciones del MMPI-2 y, crecientemente, del NEO-PI-R, y es probable que la disparidad continúe.

PRUEBAS PROYECTIVAS DE LA PERSONALIDAD

Las pruebas proyectivas representan el segundo gran enfoque sobre la evaluación de la personalidad, que es radicalmente diferente de los métodos usados en las pruebas objetivas de la personalidad. El formato, los reactivos, la aplicación y la calificación de las pruebas proyectivas de personalidad son distintos de los de las pruebas objetivas. Mientras que las pruebas objetivas requieren respuestas a preguntas o aseveraciones verbales explícitas, las pruebas proyectivas piden respuestas a estímulos ambiguos y no estructurados. De hecho, un rasgo importante que distingue a las técnicas proyectivas es el uso de tareas relativamente no estructuradas que permiten un número casi ilimitado de respuestas (Anastasi, 1997). Este rasgo significa que las instrucciones proporcionadas al examinado y los mismos estímulos tienen un nivel mínimo de estructura, lo cual se basa en la suposición de que es más probable que la dinámica interna de la personalidad del examinado se revele en circunstancias no estructuradas. Esos estímulos incluyen manchas de tinta, imágenes de naturaleza ambigua, dibujos de una persona u objetos elaborados por el examinado, y frases incompletas.

Aunque existen ejemplos anteriores de técnicas proyectivas, éstas empezaron a recibir atención de los psicólogos con la publicación en 1921 del libro Psychodiagnostik y con la prueba de manchas de tinta de Hermann Rorschach. Diecisiete años más tarde, en 1938, Henry Murray, de la Harvard University, escribió Explorations in Personality donde introdujo otra prueba proyectiva importante, la prueba de apercepción temática. Las técnicas de dibujo de la figura humana siguieron poco después, en la década de 1940 (por ejemplo, Buck, 1948; Machover, 1949). A lo largo de las siguientes seis décadas, estas investigaciones fueron sometidas a una crítica intensa debido a su endeble confiabilidad y validez; de hecho, en 1968 Walter Mischel, entonces en la Stanford University, cuestionó tanto la existencia de la personalidad, como la necesidad de contar con medidas proyectivas para este constructo "inexistente". Un hecho interesante es que esos ataques no lograron disminuir el entusiasmo de muchos profesionales por las pruebas proyectivas y su uso.

El desarrollo y uso de prácticamente todas las pruebas proyectivas de personalidad se basan en la hipótesis proyectiva, un término acuñado por Frank (1948), quien escribió que las técnicas proyectivas eran, en esencia, rayos X psicológicos. De acuerdo con la hipótesis proyectiva, cuando se enfrentan con estímulos ambiguos los examinados proyectarán en los estímulos aspectos de su personalidad en un esfuerzo por darles sentido. El examinador entonces, puede trabajar con la historia a partir de las respuestas de la persona para obtener insight (introspección) sobre las tendencias de su personalidad. El concepto de proyección proviene de la formulación de Freud (1911) de los mecanismos de defensa. Al usar el mecanismo de defensa de proyección, la gente atribuye de manera inconsciente sus propios impulsos y rasgos de

personalidad negativos a las personas que la rodean. De esta forma, se supone que las pruebas proyectivas desvían las defensas conscientes de los examinados y permiten a los clínicos tener acceso a información psicológica importante acerca de los individuos de la cual ellos mismos no están al tanto. Por último, debido a que se cree que las pruebas proyectivas son sensibles a los aspectos "inconscientes" de la personalidad, mientras que las pruebas objetivas no lo son, los defensores de las pruebas proyectivas sostienen que estos instrumentos proporcionan validez superior al uso de procedimientos más estructurados (por ejemplo, Finn, 1996; Weiner, 1999).

Lindzey (1959) clasificó las técnicas proyectivas en cinco grandes categorías. Las técnicas de asociación incluyen pruebas de manchas de tinta y de asociación de palabras. Dentro de las técnicas de construcción están las pruebas de dibujo de la figura humana y pruebas de creación de historias como el TAT. Las técnicas de completamiento incluyen pruebas de frases incompletas. Las técnicas de arreglo o selección incluyen la prueba de Szondi y la prueba de colores de Luscher. Finalmente, entre las técnicas de expresión están el juego proyectivo con muñecos, marionetas y el análisis de la escritura. Los cinco tipos principales de técnicas proyectivas en la taxonomía de Lindzey se presentan en la tabla 9.3. A continuación se estudiarán dos de los métodos de evaluación proyectiva de mayor uso y más estudiados (Piotrowski y Belter, 1999): la prueba de manchas de tinta de Rorschach y la prueba de apercepción temática.

LA PRUEBA DE MANCHAS DE TINTA DE RORSCHACH

El ejemplo prototípico de una prueba proyectiva de personalidad es la prueba de manchas de tinta de Rorschach, creado por el psiquiatra suizo Hermann Rorschach en 1921. En realidad, como afirman Hunsley y Bailey (1999), la prueba de manchas de tinta de Rorschach "tiene la dudosa distinción de ser, al mismo tiempo, el más apreciado y el más criticado de todos los instrumentos de evaluación psicológica". Rorschach experimentó con un gran número de manchas de tinta que presentó a diferentes grupos psiquiátricos y, más tarde, a diferentes grupos de personas normales (por ejemplo, estudiantes, artistas).

TABLA 9.3

Los cinco principales subtipos de técnicas proyectivas y dos ejemplos de cada subtipo

ubtipo	Ejemplos	Descripción
sociación	*Prueba de manchas de tinta de Rorschach* (Rorschach, 1921)	A los examinados se les muestran 10 manchas de ti simétricas, 5 en blanco y negro y 5 en color, y lueg les pide que digan qué les parece cada mancha de ti
	Prueba de movimientos de las manos (por ejemplo, Wagner, 1962)	A los examinados se les muestran varias imágenes manos en movimiento, y se les pide conjeturar "qué puede estar haciendo" cada mano.
onstrucción	*Prueba del dibujo de lafigura humana* (Machover, 1949)	Se pide a los examinados que dibujen una persona una hoja de papel en blanco y luego se les pide que dibujen otra persona del sexo opuesto de la primera
	Prueba de apercepción temática (Margan y Murray, 1935)	A los examinados se les muestran dibujos de situac nes sociales ambiguas y se les pide que cuenten una historia concerniente a los personajes de cada dibuj
ompletamiento	*Prueba defrases incompletas de la Washington University* (Loevinger, 1976)	Se presenta a los examinados varias frases incompl (por ejemplo, "Si mi madre ...") y se les pide que c pleten cada tema.
	Estudio de imágenes defrustración de Rosenzweig (Rosenzweig, Fleming y Clark, 1947)	A los examinados se les muestran dibujos de varias situaciones frustrantes (por ejemplo, ser salpicado intención con agua por un automóvil que pasa) y se pregunta cómo responderían verbalmente en cada s tuación
rreglo/Selección	*Prueba de Szondi* (Szondi, 1947)	A los examinados se les muestran fotografías de in duos con diferentes trastornos psiquiátricos y se les pregunta a qué pacientes prefieren más y menos
	Prueba de colores de Luscher (Luscher y Scott, 1969)	A los examinados se les pide que ordenen diferente láminas coloreadas en orden de preferencia
xpresión	*Juego proyectivo con marion;;tas* (por ejemplo, Woltmann, 1960)	Se pide a los niños que desempeñen los roles de otr individuos (por ejemplo, la madre o el padre) o de mismos usando marionetas
	Análisis de escritura (véase Beyerstein y Beyerstein, 1992, para una revisión)	Se pide a los individuos que proporcionen muestras espontáneas de su escritura

Nota: Tomado de Lilienfeld, Wood y Garb (2000).

Luego seleccionó las láminas y las características de respuesta que distinguían a cada grupo. Este proceso parece similar a un ejemplo anterior del método de clave empírica utilizado en la selección de reactivos para el MMPI, esto es, Rorschach seleccionó aquellos estímulos que distinguían mejor entre grupos de gente que se sabía que diferían en ciertas características. Sin embargo, el uso hecho por Rorschach de esta aproximación no fue empírico ya que ni obtuvo ni examinó los datos de manera cuantitativa (Anastasi, 1997). Rorschach seleccionó 15 láminas para usar en esta prueba, pero debido a que su editor carecía de los fondos para reproducir 15 manchas de tima en la primera edición de Psychodiagnostik, sólo se publicaron 10 láminas. Esas 10 láminas se usan todavía en la actualidad. Cinco de las láminas son acromáticas (sombras en gris y negro) y cinco cromáticas (láminas que incluyen color).

El uso de esas manchas de tinta como prueba psicológica se vio muy afectado por la muerte de Rorschach, a la edad de 38 años, debida a complicaciones de una apendicitis sólo un año después de que publicó su único trabajo importante en esta aproximación a las pruebas proyectivas de la personalidad. Como consecuencia de su muerte prematura, le tocó a otros extender el uso del método de Rorschach, lo que resultó en el desarrollo de varias técnicas rivales. En las décadas de 1930 y 1940 cinco psicólogos estadounidenses (Samuel Beck, Marguerite Hertz, Bruno Klopfer, Zygmunt Piotrowski y David Rapaport) crearon métodos para la aplicación y calificación del Rorschach; en la actualidad esos métodos se siguen utilizando en cierta medida. Esas cinco aproximaciones reflejaban suposiciones psicoanalíticas o psicodinámicas sobre la naturaleza de la personalidad y las formas en que la dinámica interna se refleja en las respuestas a las manchas de tinta. Sin embargo, diferían entre sí en los detalles de esas suposiciones y en la forma en que la prueba debía aplicarse y calificarse (ninguno de

los cinco sistemas usaba las mismas instrucciones para el examinado). A la postre, cada sistema desarrolló procedimientos de calificación completamente separados y cultivó grupos separados de seguidores, lo cual llevó a la fragmentación del Rorschach como una herramienta de evaluación. Más que una sola prueba, durante las décadas que siguieron a su inicio el Rorschach se convirtió en al menos cinco "pruebas" diferentes.

A esta confusión se suman los resultados de una encuesta aplicada por Exner y Exner (1972) que indica que 22% de los clínicos habían abandonado por completo la calificación y que la mayoría de los encuestados usaba sus propias reglas personalizadas de calificación, basadas en alguna combinación de los cinco sistemas. Mientras que algunos clínicos usaban el Rorschach como una tarea perceptual-cognitiva, otros la empleaban como una tarea "estímulo para la fantasía" (Erdberg, 1990). La aproximación perceptual-cognitiva al uso del Rorschach asume que la tarea del examinado es organizar su respuesta a un estímulo ambiguo. Se supone que la forma en que el individuo realiza esta tarea es representativa de la forma en que responde a los estímulos ambiguos en su vida diaria, y el contenido de la respuesta es secundario a los aspectos perceptuales o estructurales de la respuesta. En contraste, la aproximación del estímulo para la fantasía es más congruente con la hipótesis proyectiva. La persona proyecta sus estados de necesidad interna en la mancha ambigua de una manera simbó1ica. El contenido de la respuesta se interpreta como una representación simbólica de su dinámica interna. La mayor parte de los procedimientos de calificación desarrollados para el Rorschach incluyen aspectos de esos dos enfoques.

Un procedimiento estándar para la aplicación del Rorschach, es el siguiente: primero se presentan las láminas al sujeto en un orden establecido y se instruye al individuo para decir al examinador lo que ve, lo que le parece cada lámina o lo que podría ser la mancha. Se espera que el examinador registre al pie de la letra cada respuesta, aunque esta tarea ha demostrado ser muy difícil, por lo que a menudo no se sigue en la práctica clínica. Segundo, después de que se han obtenido las respuestas a las 1O láminas, el examinador regresa la primera lámina al examinado y le pide que explique en qué se basó para dar su respuesta, es decir, ¿qué la hace ver de esa manera? El examinador registra también esas respuestas, incluyendo qué parte o partes de la lámina se utilizaron para emitir la respuesta y qué rasgos de la mancha (por ejemplo, la forma, el color, la localización) fueron importantes para formar el precepto o respuesta.

Aunque los sistemas para calificar el Rorschach difieren en sus detalles, los rasgos más comunes codificados y calificados de las respuestas son la localización, los determinantes, el contenido y la popularidad. La localización se refiere a la parte de la mancha con la cual el examinado asocia cada respuesta, que puede incluir una parte o toda la mancha de tinta, es decir, si el examinado usó toda la mancha o sólo una parte de ella al formular su respuesta. Los determinantes de la respuesta ("¿qué la hizo parecer un zorro?") incluyen la forma de la mancha de tinta ("porque tiene la forma de un zorro") color ("porque tiene esos dos ojos rojos"), sombreado o variaciones de negro y gris ("porque esta parte parece un pelaje gris") y la percepción de movimiento en el estímulo ("parece que el zorro está corriendo"). La calificación del contenido de la respuesta, o lo que la persona ve, varía de un sistema de calificación a otro pero con frecuencia incluye puntuaciones por la presencia de figuras humanas, detalles humanos (partes de los seres humanos como la cabeza, las manos), figuras animales, detalles animales (partes de figuras animales), diagramas anatómicos, objetos inanimados, nubes, plantas, mapas, sangre, objetos sexuales y símbolos. Por último, la popularidad de las respuestas se califica con base en la frecuencia relativa de diferentes

respuestas entre la gente en general, es decir, la tasa base de la respuesta.
Mucha gente poco familiarizada con el Rorschach cree que el contenido de las respuestas es el aspecto más importante de la prueba, es decir, creen que las personas perturbadas o afligidas ven un contenido extraño en el Rorschach, y que esto brinda al psicólogo clínico información acerca de su diagnóstico.
Una razón por la que se minimiza la importancia del contenido en el Rorschach, así como en otras pruebas proyectivas, es lo que Chapman y Chapman (1969) denominaron correlación ilusoria. Estos investigadores encontraron que los clínicos tienden a confiar en asociaciones intuitivas (pero incorrectas) con las manchas del Rorschach al formular sus interpretaciones. Así, por ejemplo, un paciente que vio ojos en las manchas del Rorschach podría ser descrito como suspicaz o paranoide, y un paciente que vio pistolas podría ser caracterizado como una persona con tendencias violentas. Aunque esas interpretaciones parecen creíbles, no existe relación empírica entre este contenido y esas interpretaciones. Las personas encuentran difícil evitar elaborar interpretaciones sobre la base de correlaciones ilusorias, incluso cuando se les pide que tengan especial cuidado en no hacerlo, particularmente cuando están intentando procesar mucha información. En realidad, la interpretación del contenido en las pruebas proyectivas, en lugar de depender de criterios de calificación más empíricamente justificados, contribuye a darle mayor ventaja al método estadístico sobre los juicios clínicos en la predicción de la conducta.
El esfuerzo más ambicioso para proporcionar un sistema integral y psicométricamente adecuado para la calificación e interpretación del Rorschach es el trabajo de John Exner, al desarrollar el sistema general para la calificación e interpretación Rorschach (por ejemplo, Exner, 1978, 1986, 1993; Exner y Weiner, 1982). El sistema general de Exner subraya la estructura de las respuestas más que su contenido y utiliza las proporciones y combinaciones de las respuestas más que calificaciones aisladas. Exner proporcionó reglas detalladas para aplicar, calificar e interpretar el Rorschach. También ofreció un conjunto de normas de respuesta para niños y adultos de poblaciones clínicas y no clínicas, aunque se han expresado críticas a esto en el sentido de que esas normas sobreestiman el grado de desajuste en la población y que son inapropiadas para usarse con grupos minoritarios. Los resultados de los esfuerzos de Exner han encontrado revisiones mixtas. Por un lado su trabajo es aclamado por traer orden al estado caótico que había existido antes en la aplicación y calificación del Rorschach. De hecho, en 1998 el comité de asuntos profesionales de la American Psychological Association elogió a Exner por su trabajo sobre el Rorschach. Por otro lado, los críticos han afirmado que el sistema general de Exner todavía no cumple los estándares que le confieran una confiabilidad y validez adecuadas. Esos críticos sostienen que ni la confiabilidad entre calificadores ni la validez creciente del sistema general de Exner se han demostrado de manera adecuada, que las interpretaciones clínicas generadas a partir de patrones específicos de respuestas no han sido apoyadas por la investigación, y que buena parte de la investigación que pretende proporcionar la base para el sistema no está publicada ni disponible para que otros investigadores la examinen. Las críticas al sistema general de Exner han sido más mordaces en relación al índice de egocentricidad (EGOI, Egocentricity Index), el índice clínico más antiguo en este sistema de calificación. El EGOI se utiliza en el sistema general como una medición de la atención centrada en uno mismo y de narcisismo, y se supone que es útil para evaluar el trastorno narcisista de la personalidad y la depresión. Nezworski y Wood (1995) analizaron la evidencia a favor del uso del EGOI con el fin de evaluar esos trastornos y concluyeron que la escala carece de confiabilidad y validez suficientes para esos propósitos.

A pesar de los esfuerzos de Exner por dar cierto grado de estandarización en la aplicación, calificación e interpretación del Rorschach, siguen proliferando métodos alternativos al uso de la prueba. Stricker y Healey (1990) citaron varios procedimientos diferentes para calificar e interpretar el Rorschach basados en la teoría de las relaciones objetales. Blatt y sus colegas (por ejemplo, Blatt, Brenneis, Schimek y Glick, 1976; Blatt y Lerner, 1983) se inspiraron en el psicoanálisis del desarrollo y en la psicología cognitiva del desarrollo para construir el análisis del desarrollo del concepto de la escala de objeto. Usando un enfoque de contenido para analizar las respuestas al Rorschach, el sistema de Blatt evalúa siete dimensiones de las relaciones objetales (por ejemplo, motivación de acción, integración objeto-acción, naturaleza de la interacción). La confiabilidad entre calificadores cae entre 80 y 90% para las escalas. Las escalas han demostrado distinguir entre poblaciones normales y psiquiátricas y también entre grupos de diagnóstico. Un segundo procedimiento, la mutualidad de la autonomía de la escala (MOAS, Mutuality of Autonomy Scale) la elaboraron Urist y sus colegas basados en el trabajo de Kohut y de Kernberg que reflejan la psicología del self, la psicología del yo y la teoría de las relaciones objetales. El MOAS se centra en el grado de separación-individuación del individuo con respecto a otra gente en el desarrollo de su identidad o yo. En esta aproximación se califican siete escalas, incluyendo reciprocidad-mutualidad, colaboración-cooperación, y analítico-dependiente. Se reporta que la confiabilidad es adecuada; por ejemplo, Urist (1977) reportó puntuaciones de confiabilidad entre calificadores que iban de .52 a .86 para el MOAS. También reportó validez de constructo basada en correlaciones con datos autobiográficos y calificaciones que el equipo realizaba de los pacientes. En un estudio de una población infantil no clínica, Ryan, Avery y Grolnick (1985) informaron de correlaciones significativas entre las calificaciones y el funcionamiento interpersonal medido por el MOAS y también ofrecieron evidencia de la validez discriminante de esta medida. Por último, Strauss y Ryan (1987) usaron el MOAS para diferenciar anoréxicas restringidas y bulímicas del grupo control.

Con base en datos como esos, Stricker y Healey (1990) escribieron:

Considerando la amplia variedad de los enfoques revisados, ¿qué puede concluirse acerca de la evaluación de los fenómenos de las relaciones objetales? En el presente, no existe un sistema general solo o enfoque para la evaluación de los fenómenos de las relaciones objetales, ni es probable que haya uno... esto se debe a que diferentes orientaciones teóricas tienen diferentes concepciones de los fenómenos de las relaciones objetales... La carencia de un sistema individual no debe verse como un problema o debilidad de las pruebas, sino como una reflexión de la diversidad de las interpretaciones teóricas de las relaciones de constructo.

Después de más de 80 años de uso y de miles de estudios, la controversia acerca del uso del Rorschach sigue presente. Docenas de revisiones han resumido la investigación sobre el Rorschach. Mientras que algunos revisores han concluido que el Rorschach es lo suficientemente confiable y válido como para emplearse en la investigación y la práctica clínica, otros han concluido que los sistemas de calificación e interpretación disponibles en la actualidad no cumplen con los criterios mínimos de confiabilidad y validez. Otros toman una posición intermedia, sugiriendo que algunas variables del Rorschach se relacionan de manera confiable con la identificación de algunas formas de psicopatología, sobre todo con la esquizofrenia y el trastorno limítrofe de personalidad. ¿Cómo puede continuar este inquietante estado de cosas después de haber puesto tanto esfuerzo por escudriñar esta prueba? ¿Cómo pueden los expertos estar en un desacuerdo tan drástico? Parece que el desacuerdo continuo acerca del estado del Rorschach se basa en diferencias en los estándares relacionados con lo que constituye una adecuada confiabilidad y validez, en diferencias en la forma en que se usa el

Rorschach en la investigación y la manera en que se usa en la práctica clínica, y en parte, en diferencias en las orientaciones teóricas de los psicólogos investigadores y los psicólogos profesionales que usan la prueba de manchas de tinta de Rorschach en su práctica clínica.

LA PRUEBA DE APERCEPCIÓN TEMÁTICA

Aunque la prueba de apercepción temática (TAT) es una prueba proyectiva en el mismo sentido amplio en que lo es el Rorschach, presenta estímulos más estructurados y requiere respuestas verbales más complejas y organizadas (Anastasi, 1997). Se presenta al examinado una serie de dibujos y se le pide que cree una historia acerca de cada imagen. La interpretación de las respuestas, por lo general, se basa en un análisis cualitativo del contenido. Desarrollado por Henry Murray y sus colegas en la clínica psicológica de Harvard (Murray, 1938, 1943), el TAT consta de 31 láminas que contienen imágenes vagas en blanco y negro y una lámina en blanco. Murray creía que el TAT es un "método para revelar al interpretador entrenado algunas de las pulsiones, emociones, sentimientos, complejos y conflictos dominantes de la personalidad. Su valor especial está en el poder para exponer tendencias inhibidas subyacentes que el sujeto no está dispuesto a admitir o no puede admitir debido a que no está consciente de ellas". Eligió el término apercepción en lugar de percepción porque creía que los examinados interpretan de manera activa los estímulos del TAT de acuerdo con sus rasgos de personalidad y sus pulsiones.

Figura 9.3
Un ejemplo de una lámina de la prueba de apercepción temática (TAT). *(Reproducido con autorización de los editores de Henry A. Murray.* Prueba de apercepción temática, *Cambridge, Mass: Harvard University Press. Copyright 1943 por el presidente y miembros del cuerpo docente de Harvard College, 1971 por Henry A. M urray .)*

Usualmente sólo se presentan al examinado de ocho a 12 de las 31 láminas, seleccionadas de acuerdo al problema de la persona. Se pide al examinado que construya una historia que se ajuste al dibujo de cada lámina, que diga qué condujo al evento mostrado en el dibujo, que describa lo que sucede en el momento y lo que los personajes están pensando y sintiendo, y que exponga un final para esta escena. En la figura 9.3 se presenta un ejemplo de una lámina del TAT. Las siguientes son las instrucciones originales de Murray (1943) para el examinado:

Ésta es una prueba de imaginación, una forma de inteligencia. Voy a mostrarle algunos dibujos, uno a la vez y su tarea será elaborar una historia dramática para cada una. Diga qué ha llevado al suceso mostrado en el dibujo, describa lo que sucede en el momento, lo que los personajes están sintiendo y pensando; y luego dé el final. Exponga sus pensamientos conforme le vengan a la mente. ¿Entiende? Como tiene cincuenta minutos para 10 dibujos, puede dedicar alrededor de cinco minutos para cada historia .Aquí está el primer dibujo (p. 3)

El examinador registra las respuestas completas del individuo para calificarlas más tarde. Cada respuesta se interpreta determinando quién es el "héroe" de la historia, es decir, el personaje de cualquier sexo con el cual presumiblemente se identifica el examinado. El sistema original de calificación continúa luego con el análisis del contenido en términos de las 17 necesidades del

modelo de personalidad de Murray. Anastasi (1997) informa que se han acumulado datos normativos y que se han desarrollado sistemas estructurados de calificación 'que logran una confiabilidad adecuada, pero como son tan engorrosos los evaluadores utilizan con más frecuencia sus propias normas y estándares subjetivos. De hecho, la mayor parte de los clínicos también varían los métodos de aplicación, número, secuencia y tipos de láminas que proporcionan, y sólo 3% informa que utiliza algún sistema estandarizado de calificación. Como puede deducirse, semejante forma de aplicación es muy problemática porque es imposible evaluar la confiabilidad y validez de los métodos de calificación e interpretación usados por examinadores individuales en la práctica clínica. Como Ryan (1985) afirmó, "Es probable que los profesionales que interpretan el TAT usen sistemas diferentes, una combinación idiosincrásica de sistemas o ningún sistema en absoluto. Ésta es la pesadilla de los psicómetras, y también sugiere que en el uso común la interpretación del TAT se basa en estrategias de confiabilidad y validez desconocidas y no probadas, un resultado potencialmente peligroso.

Se han desarrollado varios sistemas de calificación para el TAT. Por ejemplo, Bellack (1993) evalúa cada historia en varias categorías, incluyendo estructura inconsciente y pulsiones del sujeto, relaciones con los demás, conflictos importantes, defensas usadas y fortaleza del yo. Westen y sus colegas (por ejemplo, Westen, 1991; Westen, Lohr, Silk, Gold y Kerber, 1990) crearon un sistema de calificación para el TAT diseñado para evaluar las relaciones objetales, es decir, las representaciones mentales que el examinado tiene de otras personas. Westen evalúa cuatro dimensiones de las relaciones objetales: complejidad de las representaciones de la gente, tono afectivo de las relaciones, capacidad para la inversión emocional en las relaciones y estándares morales, y comprensión de la causalidad social. Sin embargo, como Lilienfeld (2000) advierten, varios estudios que examinaron el sistema de Westen utilizaron diferentes láminas del TAT, lo que dificulta realizar comparaciones entre estudios. No obstante, este sistema puede ser útil para distinguir a los individuos que han sido diagnosticados con trastorno limítrofe de personalidad de ambos, de controles no psiquiátricos y de los individuos con otras formas de psicopatología.

La investigación más amplia sobre el TAT, encabezada por McClelland y Atkinson en las décadas de 1960 y 1970, utilizó un subconjunto de las láminas originales del TAT para medir la necesidad de logro (1938) o motivación de poder de Murray (por ejemplo, McClelland , 1985; McClelland, Atkinson, Clark y Lowell, 1976). McClelland y Atkinson definieron el logro, o motivación de poder, como la necesidad de poder y de tener impacto o de afectar a otra gente. Los individuos con una elevada motivación de logro, por lo general, son asertivos, discutidores y competitivos; tienden a coleccionar objetos de prestigio o a unirse a organizaciones para tener un mayor sentido de impacto interpersonal (McClelland). McClelland también estaba interesado en la motivación de afiliación, definida como una fuerte necesidad de ser aceptado y de estar cerca de otra gente. Los individuos que califican alto en el motivo de afiliación desean ser amistosos y agradar a los demás, no como un medio para lograr un fin, como es el caso de la gente con alta motivación de logro, sino como un fin en sí mismo. Una serie de estudios de Jemmott, McClelland y sus colegas se han centrado en la asociación de la motivación de poder con dos importantes criterios objetivos externos: la disfunción cardiovascular y la función inmune. McClelland y Jemmott (1980) encontraron que estudiantes universitarios con una elevada motivación de poder presentaban más enfermedades que los estudiantes con una elevada motivación de afiliación. Para examinar los procesos biológicos subyacentes que podrían contribuir a las tasas diferenciales de enfermedad entre las personas que diferían en su motivación de poder y de afiliación, McClelland, Alexander y Marks (1982) se concentraron en la

inmunoglobulina salival A (S-IgA), una medida de la función inmune que refleja susceptibilidad a la enfermedad. Usando una muestra de prisioneros masculinos, McClelland et al., encontraron que la mayor motivación de poder se relacionaba con menor S-IgA. Este hallazgo puede tener considerables repercusiones prácticas debido a que la S-IgA refleja la función hormonal inmune y es la primera línea de defensa del cuerpo contra muchas formas de infección.

Jemmott y sus colegas profundizaron esta investigación concentrándose en la actividad de las células asesinas naturales, una medida de la inmunidad mediada por las células. La actividad de las células asesinas naturales refleja el grado en que estas células destruyen a células objetivo. Las células asesinas naturales actúan sin una sensibilización previa y parecen ser un importante componente de la defensa del cuerpo contra muchos virus y neoplasmas. Jemmott et al., (1990) evaluaron los niveles de motivación de poder y motivación de afiliación basándose en las respuestas a cuatro imágenes del TAT en muestras de estudiantes universitarios, hombres de clase media y pacientes en una organización de cuidado de la salud. De manera importante, sus resultados indicaron que mientras la motivación de poder, evaluada por la respuesta al TAT, se asociaba con una actividad relativamente baja de las células asesinas naturales, la motivación de afiliación estaba relacionada con una mayor actividad de las células asesinas naturales.

Un estudio similar lo realizó McKay (1991), quien presentó a sus sujetos una película que activaba sentimientos acerca de las relaciones y luego midió las respuestas a imágenes similares al TAT en la escala afiliativa de confianza-desconfianza. McKay consideró que esta escala mide aspectos o cualidades de las relaciones objetales y sus asociaciones con la función inmune (concentración de S-IgA). Encontró que los sujetos que producían descripciones benevolentes en respuesta a las imágenes del TAT tenían incrementos en su función inmune después de ver la película, mientras que los que describían relaciones negativas tenían decrementos en su función inmune. En general, McKay halló que las altas calificaciones de desconfianza basadas en las respuestas al TAT se asociaban de manera significativa con un funcionamiento inmunológico pobre.

La investigación que relaciona las respuestas al TAT con los parámetros de la función inmune es importante por dos razones. Primero, esos estudios ofrecen evidencia a favor del vínculo entre los factores emocionales y la resistencia del cuerpo a la enfermedad. Esta asociación pone de relieve la probabilidad de que los factores psicológicos desempeñen una función en el inicio, el curso o la recuperación de enfermedades infecciosas. Segundo, esos estudios muestran evidencia a favor de la asociación entre las respuestas a una prueba proyectiva y criterios externos duros y objetivos. Es decir, sugieren que si una prueba proyectiva como el TAT puede calificarse de manera confiable, puede demostrar ser un predictor válido de criterios externos más claros y observables. Se requiere de más investigaciones como las realizadas por Jemmott y McKay para proporcionar una base científica clara del uso de pruebas proyectivas como un método para formular predicciones de criterios conductuales y somáticos altamente específicos.

En suma, como Lilienfeld et al. (2000) notaron, existe un apoyo modesto para la validez de constructo de una serie de temas de calificación del TAT, sobre todo los que evalúan las relaciones objetales y la necesidad de poder, logro y afiliación. Sin embargo, hay también problemas no resueltos concernientes a lo apropiado del TAT para uso clínico. Incluso con el más confiable de los sistemas de calificación del TAT, por lo general no se dispone de normas adecuadas para comparar a los examinados con los perfiles de individuos de diferentes grupos de diagnósticos o de diferentes grupos culturales. Más aún, existen otras medidas que evalúan las relaciones objetales y la motivación de logro o afiliación. Por consiguiente, no queda claro

que el TAT sea necesario para evaluar esos constructos.