

La psicología científica y el Test de Rorschach¹

Irving B. Weiner
Universidad del Sur de Florida
Charles D. Spielberger
Universidad del Sur de Florida
Norman Abeles
Universidad del Estado de Michigan

En la edición previa de la revista *The Clinical Psychologist*, Lohr, Fowler y Lilienfeld (2002) expresaron su preocupación acerca del uso de los métodos de evaluación y tratamiento que carecen de validación científica. Al describir estos métodos como “pseudociencia”, los autores se lamentan de la magnitud en que la popularidad de algunas de estas técnicas y las reivindicaciones realizadas a su favor sobrepasan las escasas evidencias científicas que los avalan. Acabaron solicitándole a la *American Psychological Association (APA)* [Asociación Psicológica Americana] y a otras organizaciones de profesionales que “se impongan severas sanciones, incluida la expulsión si fuera necesario, para los médicos que rutinariamente utilizaran prácticas terapéuticas y de evaluación carentes de aval científico” (p. 8). A lo largo de todo su artículo los autores mencionan el Test de Rorschach (TR) como ejemplo de la “ciencia chatarra” en la evaluación y aparentemente como candidata para las sanciones que solicitan.

Aprobamos absolutamente el apropiado lugar que ocupan los métodos científicos en la práctica y la investigación psicológica clínica y coincidimos con el Sr. Stricker (1997) respecto de que todos los médicos clínicos, ya sea quienes trabajen en un laboratorio o quienes brinden servicios para la salud mental, deben luchar por orientar sus tareas desde una perspectiva científica. Sin embargo, clasificar al Test de Rorschach

¹ Enviar la correspondencia relacionada con este artículo a: Irving B. Weiner, Ph.D., 13716 Halliford Drive, Tampa, FL 33624, iweiner@hsc.usf.edu

como un procedimiento pseudocientífico que debiera ser eliminado y sus profesionales expulsados de la psicología, va contra la corriente ante la abundante literatura de investigación que registra la sensatez psicométrica del instrumento. El presente artículo sintetiza los datos clave al respecto y brinda rectificaciones de diversas afirmaciones engañosas presentadas por Lohr y otros (2002).

Sensatez psicométrica del Test de Rorschach

La sensatez psicométrica de un instrumento de evaluación se define mediante procedimientos estandarizados, concordancia entre los codificadores, fiabilidad, datos normativos y validez. En el transcurso de la reciente controversia entre los defensores y los críticos del Test de Rorschach, en general se coincidió en que este instrumento, si se lo administra de manera adecuada, consiste de un estandarizado conjunto de materiales que se presenta con instrucciones uniformes ante las personas que responden. Respecto de los otros cuatro componentes psicométricos, los críticos del Rorschach plantearon diversas cuestiones pero, la gran mayoría siguió sin poder convencerse con los datos empíricos generados para resolver estas preocupaciones o con los esfuerzos que realizaron los investigadores del Rorschach para responder a las críticas mediante la realización de estudios adicionales que caracterizan al método científico.

Concordancia entre los codificadores

Los críticos del Rorschach alegaron que los datos originales que muestran una importante concordancia entre los codificadores para el Sistema Comprehensivo de Rorschach (por ejemplo, Exner y Weiner, 1995, pp. 21-27) es poco fiable porque se basan en un porcentaje de concordancia en lugar de basarse en estadísticas más sofisticadas como los coeficientes Kappa o de Correlación Intraclase que consideran la concordancia casual (Wood, Nezworski y Stejskal, 1996). Posteriormente, las revisiones metaanalíticas y los estudios con muestras de pacientes y pacientes ambulatorios

identificaron coeficientes Kappa medios a lo largo de diversas categorías de codificación del Sistema Comprehensivo que oscilaron desde .79 hasta .88, lo cual indica un excelente rango para el método Kappa (Acklin, McDowell, Verschell y Chan, 2000; Meyer, 1997a, 1997b). Respecto de las Correlaciones Intraclase, Meyer y otros (2002) observaron medianas y medias de los coeficientes interjuez de .92 y .90, respectivamente, para 164 variables de sumario estructural en dos clasificaciones independientes de 219 protocolos que incluyeron 4761 respuestas. De las 164 variables examinadas en esta investigación, 156 (95%) cumplieron con los criterios de Correlación Intraclase de la excelente fiabilidad y ninguna mostró tener poca fiabilidad. Sin duda, entonces, el TR puede codificarse fiablemente mediante el Sistema Comprehensivo.

Fiabilidad

Los críticos del Rorschach alegaron que la fiabilidad del TR aún no se ha establecido porque se han publicado nuevos análisis de las correlaciones sólo para una porción de las variables en el Sistema Comprehensivo (Garb, Wood, Nezworski, Grove y Stejskal, 2001). Como destacaron Viglione y Hilsenroth (2001), la mayoría de estos nuevos análisis de las correlaciones llamadas “ausentes” incluyen ya sea variables compuestas acerca de las cuales se dispone de datos de fiabilidad para las partes que la componen, o variables que se presentan con demasiado poca frecuencia como para permitir una significativa comparación del análisis-nuevo análisis. La verdad es que casi todas las variables codificadas mediante el Sistema Comprehensivo y concebidas como relacionadas con rasgos característicos muestran una importante estabilidad a corto y largo plazo y que los nuevos análisis de las correlaciones superan el .75. Entre éstas existen 19 importantes variables centrales que cuentan con nuevos análisis de correlaciones a 1 año o a 3 años del .85 o superiores (ver Viglione y Hilsenroth, 2001, Tabla 1). Sin duda, entonces, el Sistema Comprehensivo de Rorschach arroja datos fiables.

Datos normativos

Los críticos del Rorschach alegaron que los datos normativos de los cuales se dispone en la actualidad para el Sistema Comprehensivo, recolectados por Exner hace unos 20 años y publicados más recientemente en el año 2001 (Exner, 2001), están desactualizados, son inexactos y probablemente generen patologías exageradas sugiriendo un trastorno donde realmente no lo hay (Wood, Nezworski, Garb y Lilienfeld, 2001). Si bien Wood y otros citan observaciones de investigaciones que aparentemente avalan sus cuestionamientos, la mayoría de los estudios que mencionan presentan compromisos metodológicos como consecuencia de las reducidas y no representativas muestras de pacientes ambulatorios, de la carencia de procedimientos sistemáticos y el uso de examinadores sin experiencia (ver Weiner, 2001a). No obstante, al reconocer que los instrumentos de evaluación requieren una re-estandarización periódica, Exner (2002) asumió un nuevo proyecto de recolección de datos normativos en el cual examinadores con experiencia están analizando una muestra participante importante y demográficamente representativa con un conjunto de instrucciones uniforme. Las observaciones de las primeras 175 personas analizadas en este proyecto se asemejan mucho a los datos de referencia preliminares y dispersan los cuestionamientos acerca de la asignación de patologías exageradas. De los primeros 175 pacientes ambulatorios de la nueva muestra, sólo uno demostró un incremento en el índice del Sistema Comprehensivo del trastorno de percepción y pensamiento (PTI > 2); sólo el 16% presentó puntajes elevados en el índice de depresión (DEPI > 4) y sólo el 6% presentó índices de inhabilidad social (CDI > 3).

Profundizando respecto de los estándares normativos, los críticos del Rorschach alegaron que “las personas de raza negra, hispana, estadounidenses y no estadounidenses presentan puntajes diferentes es importantes variables del Rorschach tanto para el Sistema Comprehensivo como para otros enfoques” (Wood y Lilienfeld, 1999, p. 342). Sin informar datos del Sistema Comprehensivo específicos que avalen esta sentencia,

estos críticos concluyeron que “debido a que existen importantes diferencias transculturales y debido a que no se han desarrollado normas adecuadas, es dudoso si el Sistema Comprensivo debiera utilizarse para evaluar a los miembros de grupos estadounidenses minoritarios” (p. 341).

Sin embargo, los hechos son bastante diferentes. En primer lugar, las diferencias transculturales en los puntajes promedio del test, si existieran, podrían reflejar diferencias culturales reales que serían medidas con exactitud por estos puntajes del test. De acuerdo con esto, las diferencias obtenidas no necesariamente tienen implicancias sobre la validez de utilizar un test en culturas diversas. En segundo lugar, estudios recientes realizados con grupos estadounidenses minoritarios indican que no existen importantes diferencias en el Sistema Comprensivo entre dichas culturas. Presley, Smith, Hilsenroth y Exner (2001) observaron una diferencia clínicamente significativa en sólo 1 de 23 variables centrales seleccionadas a priori entre 44 afroamericanos y 44 estadounidenses de raza blanca demográficamente ajustados de la muestra de referencia de pacientes ambulatorios en el Sistema Comprensivo. Meyer (2002) no observó una asociación entre la etnia y ninguno de los 188 puntajes de Rorschach obtenidos entre los participantes que respondieron que incluyeron estadounidenses nativos, de ascendencia europea, hispana, asiática, afroamericanos, demográficamente ajustados, en una muestra multicultural de 432 pacientes consecutivos evaluados en un programa psicológico basado en un hospital. Mediante sofisticados procedimientos de corte/inclinación para examinar un posible sesgo en la medición del test, Meyer llegó a la conclusión de que “los datos disponibles avalan claramente el uso cross-étnico del Sistema Comprensivo” (p. 127).

Validez

Los críticos del Rorschach alegaron que el TR es un test “de mala calidad” con poco o ningún criterio o constructo de validez (Dawes, 1994; Hunsley y Bailey, 1999; Lilienfeld, Wood y Garb, 2000). Sin embargo, de hecho, en un estudio metaanalítico

realizado por Hiller, Rosenthal, Bornstein, Berry y Brunell-Neuleib (1999; ver también Rosenthal, Hiller, Bornstein, Berry y Brunell-Neuleib, 2001) de estudios de investigación del Rorschach publicados desde 1977 hasta 1997, en los cuales había al menos una variable externa (es decir, no incluida en el test) y en los cuales se había postulado alguna base razonable para las asociaciones esperadas entre las variables, se identificó un coeficiente de validez medio no ponderado del .29 para las variables del Rorschach en 2276 protocolos. Se aplicaron métodos similares en 5007 protocolos MMPI en estudios publicados durante ese mismo período que arrojaron un coeficiente de validez medio no ponderado del .30 para las variables del MMPI. Estos coeficientes de validez virtualmente equivalentes demuestran que el TR generalmente es tan válido como el MMPI. Hiller y otros (1999) concluyeron respecto de ambas mediciones que la “validez de estos instrumentos es casi tan buena como era de esperarse en los tests de la personalidad” (p. 291) y que las magnitudes del efecto tanto del TR como del MMPI justificaron el uso de estas mediciones con los fines deseados.

En 1996, el Directorio de Asuntos Profesionales de la *American Psychological Association* designó un Grupo de Trabajo de Evaluación Psicológica [*Psychological Assessment Work Group* (PAWG)] encargado de recopilar evidencias acerca de la eficacia de la evaluación en la práctica clínica, particularmente respecto de la validez y la utilidad del análisis psicológico en los servicios del cuidado de la salud. El informe final del PAWG, publicado en la revista *American Psychologist* (Meyer y otros, 2001), llegó a la conclusión a partir de su revisión de estudios predictivos y metaanalíticos de que “la validez del análisis psicológico es potente y convincente” y “comparable a la validez del análisis médico” (p. 128). Las correlaciones específicas que enumeraron con fines comparativos incluyeron las siguientes:

Peso y altura de los adultos estadounidenses (.44)

Escalas de validez del MMPI y detección de psicopatologías fingidas (.44)

CI del WAIS y nivel de educación obtenido (.44)

Puntajes de la escala para padres del Rorschach y posterior resultado de la psicoterapia (.44)

Viagra y mejoría de la función sexual masculina (.38)

Puntajes de la dependencia y la conducta dependiente en el Rorschach (.37)

Puntajes de la escala de MMPI y capacidad promedio para detectar trastornos depresivos o psicóticos (.37)

Puntajes de la Escala de Psicopatía de Hare y posterior conducta violenta (.33)

Resultados de las mamografías y detección del cáncer de mama en el término de 1 año (.32)

Pastillas para dormir y mejoría a corto plazo en el insomnio crónico (.30)

Meyer y otros (2001, p. 128) consideraron estas correlaciones como “comparables a la validez del análisis médico”.

En Viglione (1999), Viglione y Hilsenroth (2001) y Weiner (1996, 2001b) se presentan detalladas discusiones de los temas que participan en la evaluación de la validez de las variables del Rorschach y revisiones de un vasto conjunto de investigaciones que demuestran la utilidad de la evaluación de Rorschach. Para completar esta presentación de los hechos relacionados con la sensatez psicométrica del TR, se pueden extraer dos merecidos ejemplos de su constructo de validez a partir de los datos de referencia del Sistema Comprehensivo. Primero, los cambios relacionados con la edad observados en 1390 niños y adolescentes como pacientes ambulatorios de 5 a 16 años de edad, son sumamente equiparables a fenómenos de desarrollo bien establecidos en las personas jóvenes. El egocentrismo, o ego-centrismo según el concepto de Piaget (Piaget e Inhelder, 1969), se ha demostrado en la literatura de la psicología de desarrollo como un fenómeno que se reduce en función de la edad. El Índice de Egocentrismo del Sistema Comprehensivo de Rorschach, concebido como una medida del egocentrismo, presenta un nivel medio del .69 en los pacientes ambulatorios de 5 años, edad después de la cual se reduce de manera casi perfectamente lineal a .43 a los 16 años, momento en el cual

aún sigue siendo levemente más elevado que la media del .40 de un paciente ambulatorio adulto –como se podría predecir a partir de lo que se conoce acerca de las personas jóvenes (Exner, 2001, Cap. 11).

El Segundo ejemplo del constructo de validez deriva del hecho de contar con datos de referencia del Sistema Comprehensivo de 600 pacientes ambulatorios adultos y muestras de 535 pacientes psiquiátricos ambulatorios adultos, 279 pacientes internados con trastornos depresivos mayores y 328 pacientes internados con una primera internación por esquizofrenia. Teniendo en cuenta las variaciones individuales, estos cuatro grupos de adultos representan una secuencia de trastornos psicológicos cada vez más severos. Dos índices de Rorschach clave para el trastorno psicológico son el X-% (un índice del trastorno de la percepción de la realidad) y WSum6 (un índice de los trastornos del pensamiento). Si X-% y WSum6 fueran medidas válidas del trastorno, deberían incrementarse de manera lineal a lo largo de estos cuatro grupos de referencia. Consistente con esta expectativa, el valor medio de X-% aumenta de .07 en pacientes ambulatorios a .16 en pacientes ambulatorios, a .20 en pacientes internados con depresión y a .37 e pacientes internados con esquizofrenia. Los valores medios de WSum6 de estos cuatro grupos, respectivamente, fueron 4.48, 9.36, 18.36 y 42.17 (Exner, 2001, Cap. 11).

Los críticos del Rorschach a menudo parecen no notar o hacer caso omiso de la importancia de estos y otros datos que avalan la sensatez psicométrica del instrumento. Además, aquellos quienes reconocen la existencia de dichas evidencias insisten en decir que no son suficientes, que no se sienten satisfechos y que debería detenerse la enseñanza y el uso del TR hasta que se generen evidencias más convincentes de su mérito científico (por ejemplo, Garb, 1999). ¿Cuántas evidencias son suficientes y que se necesita para ser convincente? La reducida proporción de literatura de investigación relevante a la cual hacemos referencia en este artículo debiera ser más que suficiente para negar las afirmaciones de que la evaluación de Rorschach no ocupa un lugar en la psicología

clínica científica. En la edición actual de *Mental Measurements Yearbook*, Hess, Zachar y Kramer (2001) lo explican sucintamente:

“El Rorschach, utilizado con el Sistema Comprensivo, es un test de la personalidad mucho mejor de lo que sus oponentes desean reconocer” (p. 1037).

Rectificaciones de afirmaciones engañosas

Debido a que desconocen o a que prefieren ignorar la naturaleza y los objetivos buscados por el TR, Lohr y otros juzgan la utilidad de la evaluación de Rorschach con criterios inadecuados. Estos juicios inadecuados producen afirmaciones engañosas como las siguientes:

Si bien el SC [Sistema Comprensivo] de Rorschach posee cierta validez para detectar la esquizofrenia y enfermedades relacionadas, su validez para detectar la depresión, los trastornos postraumáticos del stress, la psicopatía y otras enfermedades psiquiátricas parece ser débil. Tampoco existen evidencias convincentes de que el SC sea útil para detectar el abuso sexual en niños, aún cuando se lo utilice frecuentemente para este fin (2002, p. 5).

Como primera rectificación de las implicancias engañosas de esta afirmación, cabe destacar que el TR no es una prueba de diagnóstico. Es una medida de los procesos de la personalidad. En la medida en que mide el trastorno del pensamiento, lo cual logra muy bien, ayuda en la detección de trastornos esquizofrénicos. En la medida en que mide el estado de ánimo disfórico y las cogniciones negativas, lo cual logra muy bien, ayuda en la identificación de la depresión. En la medida en que mide el malestar subjetivo, lo cual logra muy bien, ayuda en la identificación de trastornos de la ansiedad. Sin embargo, no pretende actuar como único criterio para el diagnóstico de trastornos esquizofrénicos, del estado de ánimo o relacionados con la ansiedad. De acuerdo con esto, la validez de la evaluación de Rorschach no puede y no debe medirse en función de sus correlaciones con categorías diagnósticas.

Como segunda rectificación, no debe esperarse que el TR detecte el abuso sexual en niños, ni se lo debe culpar por no hacerlo. Dada la bien establecida heterogeneidad de los estilos de la personalidad y de los patrones de reacción observados en las víctimas jóvenes del abuso, existen pocas razones para esperar que cualquier instrumento de evaluación de la personalidad identifique si un niño en particular ha sufrido abuso. Además, la aseveración de que el TR “se utiliza frecuentemente para este fin” se presenta sin evidencias y desconocemos si tales evidencias existen. Tampoco sabemos acerca de un solo caso en el cual un profesor o experto en el Rorschach haya recomendado el uso del TR para determinar si un niño ha sufrido abuso sexual. Por el contrario, dos informes de investigaciones recientes que demuestran la sensibilidad de ciertos índices del Rorschach sobre el impacto del abuso sexual llegaron a la conclusión de que estos índices “no establecen con absoluta certeza que haya habido abuso sexual” (Leavitt, 2000, p. 320) y “no debería utilizarse como único indicador de abuso sexual previo” (Kamphuis, Kugeares y Finn, 2000, p. 221). La literatura contiene abundantes advertencias similares en contra del uso de la evaluación de Rorschach con dichos fines. Las respuestas del Rorschach no indicarán si una persona tiene un problema con las bebidas alcohólicas, si es hijo único, si se crió en una granja. Lo que harán las respuestas del Rorschach es lo que el método tiene la intención de hacer – identificar las características de la personalidad del individuo. La utilidad del instrumento deriva entonces en la importancia de estas características de la personalidad identificadas para la toma de decisiones en el marco clínico, forense, cuidado de la salud, educativo y organizativo.

Además, Lohr y otros aseguran que quienes defienden a la evaluación de Rorschach han hecho declaraciones extravagantes acerca de que “este instrumento posee capacidades especiales, incluso notables” (p. 6) y “destacaron su prolongado uso y popularidad como evidencias indirectas de su mérito científico” (p. 7). Las evidencias que mencionan como declaraciones extravagantes consisten en entusiastas aprobaciones del TR en el texto de una cita del Directorio de Asuntos Profesionales de la APA [APA

Board of Professional Affairs (BPA)] realizada cuando se premió a John Exner con el Premio por los Distinguidos Aportes Profesionales al Conocimiento en 1998. Las razones para que esta cita haya sido excesivamente efusiva deben atribuírsele al BPA y no deben ser utilizados para ensuciar a toda la comunidad del Rorschach. La lectura completa de la literatura relevante revelará que los autores de los principales libros, capítulos y artículos de revistas relacionados con la Evaluación de Rorschach son casi uniformemente consistentes en fomentar una orientación conservadora para interpretar los datos del test y prevenir el exceso de los límites de lo que se debe esperar razonablemente del instrumento.

La aseveración de que los defensores del Rorschach equiparan la prolongada historia y la sostenida popularidad del instrumento con el mérito científico es igualmente engañosa. Reemplacemos un momento la retórica por la razón y permitamos que esta última nos provea de caminos confiables hacia la verdad. ¿Cuán probable es que tantos evaluadores del Rorschach hayan utilizado el instrumento durante tanto tiempo, en tantos lugares y contextos, sólo sobre la base de una correlación ilusoria? Si esto parece poco probable, ¿es irracional inferir que existió cierta utilidad en su trabajo?

Respecto de otro tema, Lohr y otros objetan que se los “ataque... ya que regularmente no utilizan el Rorschach en su propia práctica clínica o de investigación” y mencionan que “el valor científico de una técnica no puede ser evaluado sólo por sus amigos” (p. 7). Sin embargo, el tema dista mucho de ser amistoso. Este tema es interesante para las quienes evalúan el valor científico de una técnica que estén íntimamente familiarizadas con la naturaleza de la técnica y cómo funciona y para quienes no. ¿En qué campo de la ciencia son bienvenidas las críticas a los procedimientos por parte de personas que no utilizan o estudian los mismos personalmente?

Finalmente, terminamos donde empezamos, respecto de la solicitud de Lohr y otros de que se sancione a los psicólogos que realicen evaluaciones de Rorschach y de expulsarlos de la APA y otras organizaciones de profesionales. Además de haber

destacado que sus críticas del TR son sumamente injustificadas y que sus conclusiones con frecuencia son engañosas, nos preocupa que un pequeño grupo de psicólogos se tome la libertad de decidir qué es científico y qué no, y sobre esta base exigir, en efecto, la pena de muerte para la enseñanza y el uso de métodos que no pasan su inspección. Estar en desacuerdo con una creencia ampliamente sostenida o desafiar la validez de un procedimiento utilizado frecuentemente no hace que esa creencia automáticamente caiga bajo la sospecha científica o que el procedimiento sea éticamente inaceptable. Estar en desacuerdo y plantear un desafío tienen asidero en el discurso científico, pero las críticas injustificadas y la desestimación de la existencia de las que uno no comprende se parece más a la lamentable actitud de quemar los libros.

Referencias

- Acklin, M. W., McDowell, C. J., Verschell, M. S., & Chan, D. (2000). Interobserver agreement, intraobserver agreement, and the Rorschach Comprehensive System. *Journal of Personality Assessment, 74*, 15-57.
- Dawes, R. M. (1994) *House of cards: Psychology and psychotherapy built on myth*. NY: The Free Press.
- Exner, J. E., Jr. (2001). *A Rorschach workbook for the Comprehensive System* (5th ed.). Asheville, NC: Rorschach Workshops.
- Exner, J. E., Jr. (2002). A new nonpatient sample for the Rorschach Comprehensive System: A progress report. *Journal of Personality Assessment, 78*, 391-406.
- Exner, J. E., Jr., & Weiner, I. B. (1995). *The Rorschach: A comprehensive system*. Vol. 3. *Assessment of children and adolescents* (2nd ed.). New York: Wiley.
- Garb, H. N. (1999). Call for a moratorium on the use of the Rorschach Inkblot Test in clinical and forensic settings. *Assessment, 6*, 311-318).
- Garb, H. N., Wood, J. M., Nezworski, M. T., Grove, W. M., & Stejskal, W. J. (2001). Toward a resolution of the Rorschach controversy. *Psychological Assessment, 13*, 433-448.
- Hess, A. K., Zachar, P., & Kramer, J. (2001). Rorschach. In B. S. Plake & J. S. Impara (Eds.), *Fourteenth Mental Measurements Yearbook* (pp. 1033-1038). Lincoln, NE: University of Nebraska Press.
- Hiller, J. B., Rosenthal, R., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (1999). A comparative meta-analysis of Rorschach and MMPI validity. *Psychological Assessment, 11*, 278-296.
- Hunsley, J., & Bailey, J. M. (1999). The clinical utility of the Rorschach: Unfulfilled promises and an uncertain future. *Psychological Assessment, 11*, 266-277.
- Kamphuis, J. H., Kugeares, S. L., & Finn, S. E. (2000). Rorschach correlates of sexual abuse: Trauma content and aggression indices. *Journal of Personality Assessment, 75*, 212-224.
- Leavitt, F. (2000). Surviving roots of trauma: Prevalence of silent signs of sex abuse in patients who recover memories of childhood. *Journal of Personality Assessment, 74*, 311-323.
- Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest, 1*, 27-66.
- Lohr, J. M., Fowler, K. A., & Lilienfeld, S. O. (2002). The dissemination and promotion of pseudoscience in clinical psychology: The challenge to legitimate clinical science. *The Clinical Psychologist, 55*, 4-10.
- Meyer, G. J. (1997a). Assessing reliability: Critical corrections for a critical examination of the Rorschach Comprehensive System. *Psychological Assessment, 9*, 480-489.
- Meyer, G. J. (1997b). Thinking clearly about reliability: More critical corrections regarding the Rorschach Comprehensive System. *Psychological Assessment, 9*, 495-498.
- Meyer, G. J. (2002). Exploring possible ethnic differences and bias in the Rorschach Comprehensive System. *Journal of Personality Assessment, 78*, 104-129.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*, 128-165.
- Meyer, G. J., Hilsenroth, M. J., Baxter, D., Exner, J. E., Jr., Fowler, J. C., Pers, C. C., & Resnick, J. (2002). An examination of interrater reliability for scoring the Rorschach Comprehensive System in eight data sets. *Journal of Personality Assessment, 78*, 219-274.
- Piaget, J. & Inhelder, B. (1969). *The psychology of the child*. New York: Basic Books.
- Presley, G., Smith, C., Hilsenroth, M., & Exner, J. (2001). Clinical utility of the Rorschach with African Americans. *Journal of Personality Assessment, 78*, 104-129.
- Rosenthal, R., Hiller, J. B., Bornstein, R. R., Berry, D. T. R., & Brunell-Neuleib, S. (2001). Meta-analytic methods, the Rorschach, and the MMPI. *Psychological Assessment, 13*, 449-451.
- Stricker, G. (1997). Are science and practice commensurable? *American Psychologist, 52*, 442-448.
- Viglione, D. J. (1999). A review of recent research addressing the utility of the Rorschach. *Psychological Assessment, 11*, 241-265.
- Viglione, D. J., & Hilsenroth, M. J. (2001). The Rorschach: Facts, fictions, and future. *Psychological Assessment, 13*, 452-471.
- Weiner, I. B. (1996). Some observations on the validity of the Rorschach Test. *Journal of Personality Assessment, 8*, 206-213.

- Weiner, I. B. (2001a). Considerations in collecting Rorschach reference data. Journal of Personality Assessment, *77*, 122-127.
- Weiner, I. B. (2001b). Advancing the science of psychological assessment: The Rorschach Test as exemplar. Psychological Assessment, *13*, 423-432.
- Wood, J. M., Nezworski, M. T., Garb, H. N., & Lilienfeld, S. O. (2001). The misperception of psychopathology: Problems with the norms of the Comprehensive System of the Rorschach. Clinical Psychology, *8*, 350-373.
- Wood, J. M., & Lilienfeld, S. O. (1999). The Rorschach Inkblot Tests: A case of overstatement? Assessment, *6*, 341-349.
- Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996). The Comprehensive System for the Rorschach: A critical examination. Psychological Science, *7*, 3-10.