

## Introduction to the Special Series on the Utility of the Rorschach for Clinical Assessment

Gregory J. Meyer  
University of Alaska Anchorage

Psychologists have debated the clinical utility of the Rorschach for many years. In an effort to bring greater clarity to the relevant issues, a Special Series was organized for this journal. With the exception of a neutral, meta-analytic review, articles for the Special Series were solicited from scholars known to have opposing views on the Rorschach. The authors agreed to engage in a structured, sequential, and scientifically grounded dialog that focused on strengths and limitations when using the Rorschach in applied clinical settings. The debate takes place over the course of three iterations, with later articles building on and reacting to those generated earlier. This Introduction provides a rationale and overview for the full Special Series. In addition, it briefly describes the five Special Section articles published in this issue of *Psychological Assessment*. Five additional articles are expected to be published in an upcoming Special Section. In combination, these two Special Sections should provide clinicians, researchers, educators, and students with the most thorough, empirically rigorous, and up-to-date evaluation of the Rorschach's clinical utility.

For decades, psychologists have debated the value of the Rorschach as an instrument for use in clinical assessment. Historically, clinicians have embraced the Rorschach as a valuable tool for applied practice and, as a result, practitioner surveys have consistently demonstrated that the Rorschach is one of the most frequently used psychological tests (e.g., Camara, Nathan, & Puente, 1998; Watkins, Campbell, Nieberding, & Hallmark, 1995). Simultaneously, prominent scientists (e.g., Dawes, 1994; Eysenck, 1959; Jensen, 1965) have regularly criticized the psychometric foundation of the Rorschach. The spirit of this criticism can be succinctly summarized by two quotes from Jensen's 1965 review of the Rorschach in *The Sixth Mental Measurement Yearbook*. The opening page of the review states: "Put frankly, the consensus of qualified judgment is that the Rorschach is a very poor test and has no practical worth for any of the purposes for which it is recommended by its devotees" (p. 501). After reviewing a substantial portion of the existing literature, Jensen concluded: "The rate of scientific progress in clinical psychology might well be measured by the speed and thoroughness with which it gets over the Rorschach" (p. 509). More recently, Dawes (1994) advanced similar conclusions. Even though Jensen was unable to contribute to this Special Series, it would not be surprising if he concluded clinical psychology has not progressed much in the last 35 years given how the Rorschach has maintained its popularity among practicing clinicians.

Such pessimism and criticisms are countered by equally strong assertions that support the Rorschach's viability and evidence base. For instance, Weiner (1996) insisted "those who currently believe the Rorschach is an unscientific or unsound test with limited utility have not read the relevant literature of the last 20 years; or, having read it, they have not grasped its meaning" (p. 206). In support of Weiner's statement, one could turn to several recently published meta-analyses that have addressed the relationship between selected Rorschach scores and focused criterion measures, including physical illness (Bornstein, 1998), observed behavior (Bornstein, 1999), or subsequent outcome (Meyer & Handler, 1997). In these three meta-analyses, the magnitude of the validity coefficients (i.e., effect sizes) for Rorschach predictor scores were as strong or stronger than the effect sizes for other relevant predictor tests.

Given that such diametrically opposing views are proffered by respectable scholars, nonexpert consumers of this academic literature (e.g., students, faculty debating which courses to keep in a crowded curriculum, attorneys) should appropriately wonder where the truth resides regarding the Rorschach's utility. Short of thoroughly delving into the literature themselves, nonexpert readers are faced with one of three primary options for making a determination. First, one could decide not to decide, concluding the literature must be fragmented and contradictory, which makes it difficult to draw any firm conclusions. However, this option seems untenable because the Rorschach is one of the most frequently researched tests in applied clinical practice, and the vast literature on this instrument should provide greater clarity. Second, one could opt to place faith in those experts who espouse beliefs that are most congenial to preexisting sentiments about the Rorschach. On the basis of informal observation of reactions to the recent debate in *Psychological Science* on the Rorschach Comprehensive System (Exner, 1996; Wood, Nezworski, & Stejskal, 1996a, 1996b), my impression is that a number of psychologists without expertise in the literature have drawn global pro or con

---

*Editor's Note.* Stephen N. Haynes served as action editor for this article.

---

Correspondence concerning this article should be addressed to Gregory J. Meyer, Department of Psychology, University of Alaska Anchorage, 3211 Providence Drive, Anchorage, Alaska 99508. Electronic mail may be sent to afgjm@uaa.alaska.edu.

conclusions based on the latter (i.e., a priori assumptions about what ought to be true).

A third option is to search for synthesis between the opposing positions. Particularly, when faced with decades of polarized and often heated debate, it seems most likely that “truth” resides someplace between the poles that anchor the arguments. That is, the Rorschach is probably not a valid test for any and every intended purpose nor an invalid test for any and every intended purpose. The important questions then are conditional: For what purposes is the Rorschach a useful clinical instrument? For what purposes does it provide unique clinical information? Furthermore, it is reasonable to assume that the experts who have been making arguments for or against the Rorschach have at heart an interest in furthering the science and practice of clinical psychology. As such, their disagreements may largely reflect differential awareness of the literature, distinct criteria for what constitutes reasonable evidence, different thresholds and propensities to posit confounds in the existing literature, and so forth. Recognizing these differences should prove to be quite valuable in any quest to understand the Rorschach’s merits and limitations.

Thus, as a scientifically grounded discipline, clinical psychology would be served most optimally by pursuing the third option. Regardless of how pleasant or unpleasant the findings may be, synthesized and differentiated conclusions about the Rorschach’s clinical utility should be drawn from a fair review of the best available evidence. This Special Series, *The Utility of the Rorschach for Clinical Assessment*, was initiated as a structured means for achieving this goal. It was designed as a sequential, scientific dialog between a selected group of scholars known to have opposing views on the value of the Rorschach. The guiding assumption was that a sequential, scientifically grounded, point-counterpoint dialogue taking place in one journal at a fixed point in time using a core body of evidence would focus the arguments from all contributors. In doing so, the series should help reveal those issues surrounding the Rorschach that achieve clear resolution (whether pro or con), those issues that clearly still need further investigation, and those issues in the existing literature that pull for supportive interpretations by advocates and dismissive interpretations by critics (which should provide insight into instances when different criteria are used to evaluate the same evidence). Ultimately, a structured debate taking place over a core segment of the literature should reveal any unique strengths for the Rorschach as well as conditions (e.g., assessed constructs, settings, collateral assessment methods) when it is likely to be a useful clinical instrument. Simultaneously, the debate should also reveal any unique limitations to the method as well as conditions when it is unlikely to serve useful clinical goals.

The dialog for this Special Series takes place in three parts to be published in two Special Sections. Within each part, contributions come from authors known to hold a generally favorable view of the evidence as well as from authors known to hold a generally unfavorable view of the evidence. Articles that were prepared for Part 1 were sent to the authors contributing to Part 2, and articles prepared for Parts 1 and 2 will be sent to the authors contributing to Part 3. Given that later contributions to this series on the Rorschach build upon the earlier contributions, the format provides a structured opportunity for the contributors to discuss and debate the evidence presented by opposing scholars.

This format of sequential, evidence-based dialog should provide interested readers with the most thorough, empirically rigorous, and up-to-date evaluation of the Rorschach’s utility as a clinical instrument. By mapping out the Rorschach’s strengths and limitations for clinical assessment, this Special Series is expected to be an invaluable set of core readings that will inform both the science and practice of psychology. Clinicians, educators, students, and other interested parties will be oriented to the empirical literature and to the overarching methodological and theoretical considerations that are central to understanding this literature.

In an effort to make this series work, all contributors were held to three key “ground rules.” First, all were expected to present ideas, conclusions, and points of view that were firmly grounded in empirical data. If there were critical issues that had not been subject to empirical scrutiny, then authors were asked to document the absence of this relevant data. If there were theoretical or methodological limitations that adversely affected the existing empirical evidence, then authors were asked to document these issues and were encouraged to offer testable propositions that would correct these limitations.

Second, all contributors were expected to focus on issues relevant to clinical practice. Stand-alone evidence that some Rorschach scale was associated with gender differences, could predict errors in time estimation, or could predict learning a list of words was not considered pertinent evidence—unless there was a *cogent* link between these facts and some relevant clinical issue. The more meaningful clinical applications that authors were asked to focus on included such things as the contribution of Rorschach data to personality description, treatment planning, treatment processes, differential diagnosis, prediction of outcome, and measurement of change. Authors were also asked to address various psychometric, theoretical, and practical issues, including the sources of random and systematic error that affect Rorschach scores; evidence for or against incremental validity; the relationship between nomothetic research on the Rorschach and the idiographic decisions that have to be made in applied clinical practice; relevant models of personality and of assessment methodology that bear on the Rorschach; appropriate or needed validation criteria for Rorschach research; the cost–benefit considerations that emerge when considering the Rorschach in applied practice; and potential tensions that may exist between cost–benefit considerations and the scientific or clinical need to understand personality in its full complexity.

Third, the authors who were designated as the primary advocates and critics for each Part of the Rorschach Special Series were expected to be familiar with a core subset of the Rorschach literature. The core literature was operationally defined as articles that were published from 1977–1997 (the 20-year period before the series began) in five journals that regularly disseminate assessment research. These five journals were as follows: *Assessment*, *Journal of Clinical Psychology*, *Journal of Consulting and Clinical Psychology*, *Journal of Personality Assessment*, and *Psychological Assessment*. To ensure that the primary advocates and critics were familiar with this literature, the authors received a disk containing 445 abstracts that had been identified through PsycLIT as the Rorschach articles published in these five key journals during the 20-year target period. The 445 abstracts did not exhaustively cover the relevant literature because appropriate studies did not always have the term *Rorschach* listed in the PsycLIT title, abstract, or keywords. Further, these 445 abstracts only reflected a

portion of the recent literature, representing just 25% of the 1,754 Rorschach articles that PsycLIT identified for this time period. Although the primary advocates and critics were asked to focus on and emphasize the core literature defined above, they were not restricted to it. Rather, the 20 years of research in these five journals were designed to serve as a common, core base of evidence that all authors would draw upon.

Essentially, each contributor to this Special Series had two roles. First, as advocates or critics, they were expected to "carry the flag" for other scientists and practitioners who had similarly positive or skeptical viewpoints. Second, as scholars, they were expected to present ideas, conclusions, and points of view that were firmly grounded in empirical data. Because these two roles do not necessarily converge in a synergistic way, to maximize the quality of the dialog in the Rorschach Series, my editorial role was one of trying to ensure that all contributions served both goals simultaneously. That is, my goal was to help each set of authors advance their empirically grounded arguments while helping them avoid positions that may be seen as unsupported or biased. In particular, I hoped to ensure that all contributions would avoid the possibility that interested parties on the other side of the fence would simply dismiss their contribution as partisan rhetoric emerging from authors who were unfamiliar with the relevant evidence.

Although it is reassuring that some of my personal research has been cited to advance arguments by advocates and by critics (e.g., Exner, 1996; Wood et al., 1996a, 1996b), if the key question is whether the Rorschach can provide useful information, then I am clearly aligned with the advocate side of the debate. This could raise questions about my ability to fairly provide sound editorial recommendations to all contributors. To lessen the prospect of bias, in my feedback letters to authors I acknowledged how I was unlikely to be seen as a neutral party to the debate and asked to be notified if any of my recommendations seemed to counter their ability to take an empirically guided stance to the literature in a way that was consistent with their goals. Ultimately, however, whether I achieved a fair editorial balance is probably best decided by the contributors and by those who read the articles in the *Special Series*.

A brief overview of the Rorschach Special Series is provided below. Although somewhat complicated, the Special Series contains three parts that will be published in two Special Sections. The first Special Section is published in this issue of the journal, and it contains five articles written for Part 1 of the Special Series. In addition to the articles by the primary advocates and critics, Part 1 contains several ancillary articles that contribute pertinent theoretical considerations and empirical data. Although the next Special Section is still in preparation, it will contain five articles prepared for Parts 2 and 3 of this Special Series and is planned to be published within the next year.

## First Special Section on the Utility of the Rorschach for Clinical Assessment

### *Contextual Considerations*

Because any sound, scientific consideration of the Rorschach's utility for clinical decision making must be done within a framework that contains assumptions about personality organization, differences between assessment methods, how clinical assessments

are actually conducted in the trenches, how judgments can and cannot be made at various stages in the assessment process, and so on, it is optimal to explicitly recognize some of these broader assumptions and their potential implications for how the Rorschach is viewed in the clinical assessment process. As such, before moving into a focused examination of the Rorschach database, this Special Section was designed to begin with two empirically grounded articles addressing some of these general issues.

The authors of these two articles were asked to focus on the following questions (and other questions they may have deemed to be important): (a) How relevant is nomothetic data on individual test scales to the idiographic assessment decisions that must be made in clinical practice?<sup>1</sup> (b) What models of personality warrant or contraindicate the consideration of Rorschach data? (c) What models of assessment methodology warrant or contraindicate the consideration of Rorschach data? (d) What evidence exists to indicate that different assessment methods provide unique versus redundant information? (e) Assuming that assessment methods differ in the kind of information they provide, what model or models should clinicians use when they are attempting to integrate heteromethod data (i.e., how should it all fit together)? (f) What conceptual skills and factual information are required to derive accurate judgments from the Rorschach (e.g., knowledge of psychopathology, psychometrics, cognitive biases, etc.)? (g) Are there criteria that are more or less appropriate for Rorschach validation research? and (h) Assuming there are tensions between the practical values of expedience found in today's managed-care environment and the theoretical values of science that seek to understand personality in its full complexity, how do these factors impact the clinical value of the Rorschach?

These issues are addressed in the first article by George Stricker and Jerry Gold (1999). Stricker and Gold argue that distinct methods of clinical assessment have unique strengths and limitations. Further, they propose that the Rorschach is a valuable instrument for clinicians (and researchers) when the goal is to understand a person in his or her full complexity, particularly when clinicians rely on a sophisticated, theoretically driven synthesis of Rorschach data in conjunction with other sources of information.

In an effort to maintain the balance desired for this series, a second set of authors known to have critical views of the Rorschach had been enlisted to write an article for this subsection. Unfortunately, the authors decided not to submit a manuscript revision that was responsive to the peer review process. They felt the series would have more heuristic value if it emphasized diverse opinions over "nitty gritty factual correctness." Because the initial version of their manuscript had many fine qualities and raised many unique considerations not found in other contributions, and because I strongly wished to maintain the balance of pro and con articles in the series, I considered the authors' suggestion. How-

<sup>1</sup> In hindsight, this question was not stated optimally in my letter to contributors. I was not trying to suggest that nomothetic data somehow may be irrelevant to clinical decision making. Instead, my intent was to have authors consider the assumptions and pitfalls that emerge when one tries to move from the relationships between isolated test scales and criterion measures that hold across a large number of people to the context-dependent understanding of a unique person who is being evaluated with many scales derived from multiple sources of data.

ever, after discussing the matter with the Journal Editor, Stephen N. Haynes, we decided it would be antithetical to the goals of the series to consider the manuscript further.

### *An Initial Round of Articles From the Designated Advocates and Critics*

These articles were expected to draw upon the literature in order to present focused arguments regarding the strengths and limitations of the Rorschach in clinical assessment. Donald J. Viglione (1999) has written an article from the advocate position, whereas John Hunsley and J. Michael Bailey (1999) have written an article from the critic perspective.

Viglione (1999) systematically reviewed a large number of empirical Rorschach studies, with an emphasis on longitudinal outcomes, behavioral criterion measures, and evidence for the incremental validity of the Rorschach over self-report tests and interviews. In general, he argues that the Rorschach is best viewed as a behavioral task of problem solving. Furthermore, although recognizing some limitations, he asserts that the evidence demonstrates the Rorschach provides unique, incrementally useful information for informing clinical practice, particularly when the behavior sampled by the Rorschach score resembles the clinical behavior to be predicted.

Hunsley and Bailey (1999) proposed three criteria for assessing clinical utility. Although they recognize that the Rorschach is commonly accepted and used in clinical practice (their first criterion), they assert that the evidence, on the whole, does not support the psychometric validity of the Rorschach (their second criterion). With respect to their third criterion, even though they acknowledge that evidence is lacking for all personality tests (also see Meyer et al., 1998), they point out how research has not determined whether using the Rorschach leads to better clinical decisions for patients. Ultimately, they conclude there is no scientific basis for justifying use of the Rorschach in clinical practice.

### *A Broad Meta-Analysis*

The next article in this Special Section was designed to be an impartial meta-analysis. This review is presented in the article by Jordan B. Hiller, Robert Rosenthal, Robert F. Bornstein, David T. R. Berry, and Sherrie Brunell-Neuleib (1999). Initially, these authors were asked to summarize the following two key pieces of Rorschach validity data: (a) hypothesized *heteromethod* criterion-related validity coefficients and (b) hypothesized *monomethod* discriminant validity coefficients. Although these two types of coefficients provide the most stringent and important criteria for test validity, a preliminary survey suggested that very few studies presented clear hypotheses or data related to monomethod discriminant validity. Thus, Hiller et al. (1999) only focused on hypothesized criterion related validity.

Although this Special Section is focused on the Rorschach, scientists and clinicians need relevant benchmarks to make sense out of data and to make accurate judgments. Therefore, Hiller et al. (1999) also completed a second meta-analysis focused on the MMPI. Having one meta-analysis that examines the general validity of the Rorschach and another that examines the general validity of the MMPI allows for a meaningful comparison between these two very different assessment procedures. The research by

Hiller et al. also provides a psychometrically sophisticated update to Parker, Hanson, and Hunsley's (1988) widely cited meta-analysis. Ultimately, like Parker et al., Hiller et al. concluded that the Rorschach and MMPI produce essentially equivalent validity coefficients.

### *Focused Recommendations for Assessing the Incremental Validity of Rorschach Scores*

The final article in Part 1 of the Special Series was written by Robyn M. Dawes. Dawes (1999) provides principles that guide researchers to examine the extent of novel information (i.e., incremental validity) contained in a Rorschach score. His recommendations apply equally well to instances when one wishes to evaluate a new Rorschach score relative to existing scores or when one wishes to evaluate the unique contribution of the Rorschach relative to some other source of information. Dawes exemplifies these issues by drawing upon two data sets. Although the emphasis of his article is on the principles for conducting incremental validity analyses, one can also draw validity inferences from the data he analyzed. Given this, it is worth noting the origin of these data sets. I offered to try obtaining any data set or sets Dawes wished to examine. Because his interest was in studies that used a reasonable, clinically relevant criterion measure, he selected the two data sets described in his article. These were the only two data sets requested.

### *Second Special Section on the Utility of the Rorschach for Clinical Assessment*

#### *A Second Round of Articles From New Advocates and Critics*

After reviewing the previous five articles, the authors for Part 2 of the Special Series will write responses to issues raised in any of these articles. Their task is to debate, dispute, and question the conclusions drawn in the this Special Section's articles, to identify any points of agreement, and to raise new issues as they see fit. Irving B. Weiner will be writing from an advocate position. Howard N. Garb, James M. Wood, M. Teresa Nezworski, William M. Grove, and William J. Stejskal will be writing from the perspective of critics.

#### *A Final Round of Articles From the Initial Advocates and Critics*

After reviewing all the previous articles, the authors of the initial pro and con articles written for Part 1 of the Special Series will provide final comments and recommendations for the next generation of Rorschach research. Donald J. Viglione will write from the position of an advocate, whereas John Hunsley and J. Michael Bailey will write from a critical perspective.

For the final article in this Special Series, I will briefly summarize and comment on some of the more important issues that emerged. In particular, I hope to identify those issues that achieve a clear consensus from the existing evidence and underscore implications or directions for the future.

### *Closing Comments*

As a parting note, I offer thanks to a number of individuals who have advanced this Special Section. First and foremost, I would

like to extend my sincere gratitude to all of the contributing authors. I believe each article provides a sound contribution that is based on thoughtful deliberation and conscientious attention to the empirical literature. Second, each article in this Special Section received expert input from 3 reviewers, all of whom provided insightful and helpful commentary. The reviewers who provided input for this Special Section are as follows: Marvin W. Acklin, Robert P. Archer, Yossef S. Ben-Porath, George J. DuPaul, Philip S. Erdberg, John E. Exner, Jr., Stephen E. Finn, Ronald J. Ganellen, Mark Hilsenroth, Blair T. Johnson, Nancy Kaser-Boyd, Bill N. Kinder, Radhika Krishnamurthy, Kevin Parker, Bruce L. Smith, Donald K. Routh, Auke Tellegen, and Nathan Weed. Finally, I would like to thank Stephen N. Haynes. Not only did he envision the practical value of a Special Section that rigorously reviews the Rorschach, but his impeccable commitment to good science has provided me with an outstanding role model and very practical advice as I aspire to fill my role in this and the upcoming Special Section on the Rorschach.

### References

- Bornstein, R. F. (1998). Interpersonal dependency and physical illness: A meta-analytic review of retrospective and prospective studies. *Journal of Research in Personality, 32*, 480–497.
- Bornstein, R. F. (1999). Criterion validity of objective and projective dependency tests: A meta-analytic assessment of behavioral prediction. *Psychological Assessment, 11*, 48–57.
- Camara, W., Nathan, J., & Puente, A. (1998). *Psychological test usage in professional psychology: Report to the APA Practice and Science Directorates*. Washington, DC: American Psychological Association.
- Dawes, R. M. (1994). *House of cards: Psychology and psychotherapy built on myth*. New York: Free Press.
- Dawes, R. M. (1999). Two methods for studying the incremental validity of a Rorschach variable. *Psychological Assessment, 11*, 297–302.
- Exner, J. E., Jr. (1996). A comment on "The Comprehensive System for the Rorschach: A critical examination." *Psychological Science, 7*, 11–13.
- Eysenck, H. J. (1959). Review of the Rorschach Inkblot Test. In O. K. Buros (Ed.), *The fifth mental measurements yearbook* (pp. 276–278). Highland Park, NJ: Gryphon 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.
- Jensen, A. R. (1965). Review of the Rorschach Inkblot Test. In O. K. Buros (Ed.), *The sixth mental measurements yearbook* (pp. 501–509). Highland Park, NJ: Gryphon Press.
- Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Kubiszyn, T. W., Moreland, K. L., Eisman, E. J., & Dies, R. R. (1998). *Benefits and costs of psychological assessment in healthcare delivery: Report of the Board of Professional Affairs Psychological Assessment Work Group, Part I*. Washington, DC: American Psychological Association.
- Meyer, G. J., & Handler, L. (1997). The ability of the Rorschach to predict subsequent outcome: A meta-analysis of the Rorschach Prognostic Rating Scale. *Journal of Personality Assessment, 69*, 1–38.
- Parker, K. C. H., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin, 103*, 367–373.
- Stricker, G., & Gold, J. R. (1999). The Rorschach: Toward a nomothetically based, ideographically applicable configurational model. *Psychological Assessment, 11*, 240–250.
- Viglione, D. J. (1999). A review of recent research addressing the utility of the Rorschach. *Psychological Assessment, 11*, 251–265.
- Watkins, C. E., Jr., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26*, 54–60.
- Weiner, I. B. (1996). Some observations on the validity of the Rorschach inkblot method. *Psychological Assessment, 8*, 206–213.
- Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996a). The Comprehensive System for the Rorschach: A critical examination. *Psychological Science, 7*, 3–10.
- Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996b). Thinking critically about the Comprehensive System for the Rorschach: A reply to Exner. *Psychological Science, 7*, 14–17.

Received April 19, 1999

Accepted April 29, 1999 ■