International Journal of Personality Psychology
2017, Vol. 3, No. 1, 36-49

What Rorschach performance can add to assessing and understanding personality

Gregory J. Meyer

Department of Psychology, University of Toledo, USA

I provide a contemporary overview of Hermann Rorschach's inkblot task, including how it was developed, why it seems to work as it does, and how to contextualize inferences drawn from the things that people see, say, and do while completing the task. Following this, I review the meta-analyses that have been conducted concerning Rorschach validity, several multi-sample studies concerning focused topics, and a selection of recent individual studies. The aim of this part of the article is to illustrate the validity of the Rorschach as a behavioral performance task that can provide a useful complement to self-reported characteristics – both in clinical practice and in research on personality processes. Administering and coding Rorschach's task is much more time consuming than many other sources of personality data, particularly the ubiquitous introspective self-report method. However, it is argued that Rorschach performance provides a unique source of information about people that can validly add to the assessment and understanding of personality and psychological processes. As such, despite its history of controversy, it is an instrument that is worth consideration or reconsideration by personality assessors and researchers.

Keywords: Rorschach inkblots, personality, performance assessment, research review

For much of the history of the Rorschach, rather polarized claims about its merits have been made on both sides of the argument (Searls, 2017). Over the last 20 years, starting in 1995, the Rorschach was the subject of heated criticisms in the literature (Lilienfeld, Wood, & Garb, 2000; Nezworski & Wood, 1995; Wood, Lilienfeld, Garb, & Nezworski, 2000; Wood, Nezworski, Lilienfeld, & Garb, 2003; Wood, Nezworski, & Stejskal, 1996). The main critiques focused on Exner's (1974; 1996; 2003) Comprehensive System, though they also have encompassed the Rorschach in general. These critiques led to back and forth debates focused on reliability, validity, utility, and incremental validity (for a sequential, structured debate see Meyer, 1999; 2001b; Meyer & Archer, 2001); with the latter addressing what the Rorschach provides that cannot be obtained from another less expensive source of information (Hunsley & Meyer, 2003).

Many psychologists dismiss the Rorschach as an antiquated measure. It is not uncommon for those who use or conduct research on the Rorschach to hear people say things like, "I didn't realize that people still used that!" In this context, an important question is why personality psychologists would want to bother with the Rorschach? I believe there are good reasons why Rorschach assessment is still in the clinical psychology training curriculum and used in applied clinical and forensic practice. These same reasons provide an argument for why the Rorschach can be usefully included in research on personality and psychological processes more broadly.

Correspondence to: Gregory J. Meyer, Department of Psychology, University of Toledo, Mail Stop 941, 2801 W. Bancroft St., Toledo, OH 43606. E-mail: Gregory.Meyer@utoledo.edu

I have two main goals in this article. One is to review what I consider to be the contextual foundation for Rorschach based assessment, which includes what seemed to be Hermann Rorschach's intent when developing the inkblots. Following that I will provide a review of the meta-analytic research that has been completed on the Rorschach, and then a selective overview of some Rorschach studies, with the aim of illustrating both how it operates and how it could be applied in personality research.

Why the Rorschach?

In thinking about the question of what Rorschach assessment can add to the science of personality psychology, it is useful to note that, at least in the English language literature, psychology increasingly is becoming what Baumeister, Vohs, and Funder (2007) have called the "science of self-reports and finger movements." Self-reported methodology and finger movements in reaction time tasks are the primary sources of information that fuel the research that is being done in clinical psychology, personality psychology, social psychology, and cognitive psychology. Baumeister et al. (2007) examined how frequently actual behavior was being studied across decades in the *Journal of Personality and Social Psychology* from the 1960s through the first decade of the 2000s. In the 60s, about half of all studies examined behavior. In the 70s this jumped up to about 80%, though it has been in steady decline since that time. In the 2000s, behavior was studied in just about 15% of published articles. As a discipline, mainstream psychology is no longer studying what people do; it is studying what they report and how quickly they react.

Introspective self-report is an important way of understanding people, but it is fundamentally limited, as is every assessment method. The data indicate that when targeting seemingly parallel constructs, self-assessed characteristics are modestly to moderately associated (r = .20 to .45) with the reports of knowledgeable others (e.g., Achenbach, Krukowski, Dumenci, & Ivanova, 2005; Connelly & Ones, 2010; De Los Reyes et al., 2015; Klonsky, Oltmanns, & Turkheimer, 2002; Meyer et al., 2001; Meyer, 2002), minimally to modestly (r = .00 to .20) associated with measures of maximal performance (Beaudoin & Desrichard, 2011; Buchanan, 2016; Cyders & Coskunpinar, 2011; Meyer et al., 2001) or typical performance (Mihura, Meyer, Dumitrascu, & Bombel, 2013; Spangler, 1992), and often surprisingly uncorrelated with official records of recalled experiences (e.g., Hardt & Rutter, 2004; Henry, Moffitt, Caspi, Langley, & Silva, 1994). Thus, there are many instances when self-report will be insufficient to address applied needs, and using valid scales from assessment methods other than self-report can provide incrementally valid data to inform an assessment and to meaningfully address complex or sophisticated referral questions.

In the foregoing, maximal performance measures were differentiated from typical performance measures (see Cronbach, 1990). Maximal performance measures are tasks where people are told to perform their best, given clear task instructions, and educated about how to solve the problem. Under these conditions, they put forth their maximum effort at doing so. This encompasses intelligence tests, memory tests, tests of academic achievement, and so forth. Maximal performance tasks contrast with typical performance measures. Typical performance measures do not impose strong demands on the person. Rather they provide wide latitude for responding and leave it up to the person being assessed to decide how to go about completing the task they were asked to engage in. The Rorschach is a performance task, but it is a typical performance measure, not a maximal performance measure.

The correlations between self-reported characteristics and either maximum performance characteristics or typical performance characteristics fall in the range from zero up to about .20. They are much more likely to be in the range near zero. What this means is that when valid scales from performance methods, either maximal or typical, are used in personality assessment or personality research, they provide additional, incrementally valid data that can inform an assessment and meaningfully address psychological functioning in a manner that complements self-reported characteristics.

But why the Rorschach inkblot task in particular? The main reason is because it is a problem solving task based on perception. The task provides a standardized, in vivo sample of how people actually construe and see things (Searls, 2017). Even though Hermann Rorschach called his inkblots "accidental forms," they were anything but. Instead, they were artistically created and enhanced; Rorschach carefully selected them, pilot-tested them, and then modified them so that each one would more clearly do what he intended. What he intended, it appears, is to develop stimuli that contained multiple suggestive but incomplete or imperfect perceptual likenesses that formed competing visual images. He developed stimuli that were highly evocative to most people, as well as structured in two ways: (a) to provide some suggestive and rather obvious shapes that many people would see, and (b) to provide perceptual "hooks" that would capture or trigger personalized and unique imagery. Stated differently, it seems what Rorschach devised was stimuli with two features: the inkblots pretty readily seem like something to most people, though what is seen diverges into fairly common shared images and very uncommon uniquely seen images. These qualities make it easy for people to engage with the task and see something other than just an inkblot – which is not easy to do. Furthermore, the things that are seen in the inkblots vary as a function of the perceiver; most people see things that other people do not.

Rorschach was trained as an artist and he was fascinated with contradictory or multiply suggestive images. At the Rorschach Archives (see references) in Bern Switzerland, there are many examples of his artwork, which showed he was fascinated with form, color, and movement, as well as with contradictory images, shaped like one thing but colored like another, for example. He was very accomplished as an artist, and he put that artistic expertise to work in the inkblots.

As mentioned already, Rorschach artistically embellished and enhanced his inkblots. It appears that these images evolved with practice, with him using them, refining them, and using them again (Searls, 2017). At the Archives there are multiple versions of each of the 10 final cards. These alternatives vary in their composition and emphasis, though the same fundamental structure persists, such that it is fairly easy to recognize variants of the same card. A good example can be found in Searls (2017).

Thus, it seems that even though Rorschach did not write about this process, his intent was to design a task that was an empirical assessment of personality engaged in a particular problem solving action. He designed the inkblots with this aim; to make the images look like one thing in one part of the inkblot, but look like something else contradictory in another part of the inkblot. Providing these contradictory "critical bits," or stimulus cues, provides a stimulus situation that helps to both assess the conventionality of perception and to maximize individualized unique expression. As such, people attend to the inkblots in very distinctive ways given their psychological makeup and characteristics.

This also appears to be what makes Rorschach's inkblots more useful than other inkblots that have been developed. The inkblots that Rorschach prepared have both more internal structure and more latitude for idiographic richness than other sets of inkblots, like the Behn inkblots, the Holtzman inkblots, the Somatic Inkblot Series, or the Zulliger inkblots, none of which have reached or sustained the popularity of Rorschach's inkblots.

Interestingly, with respect to Rorschach's apparent intentions, the task creates what is known as a Zipf distribution (or Pareto or power-law distribution) of objects perceived. Zipf (see references) was a linguist who discovered how the frequency of words used in a language were lawfully related to their rank, with the first rank being the most common word, the second rank being the second most common word, and so on. The second most common word, which in English is the word "of," occurs about half as

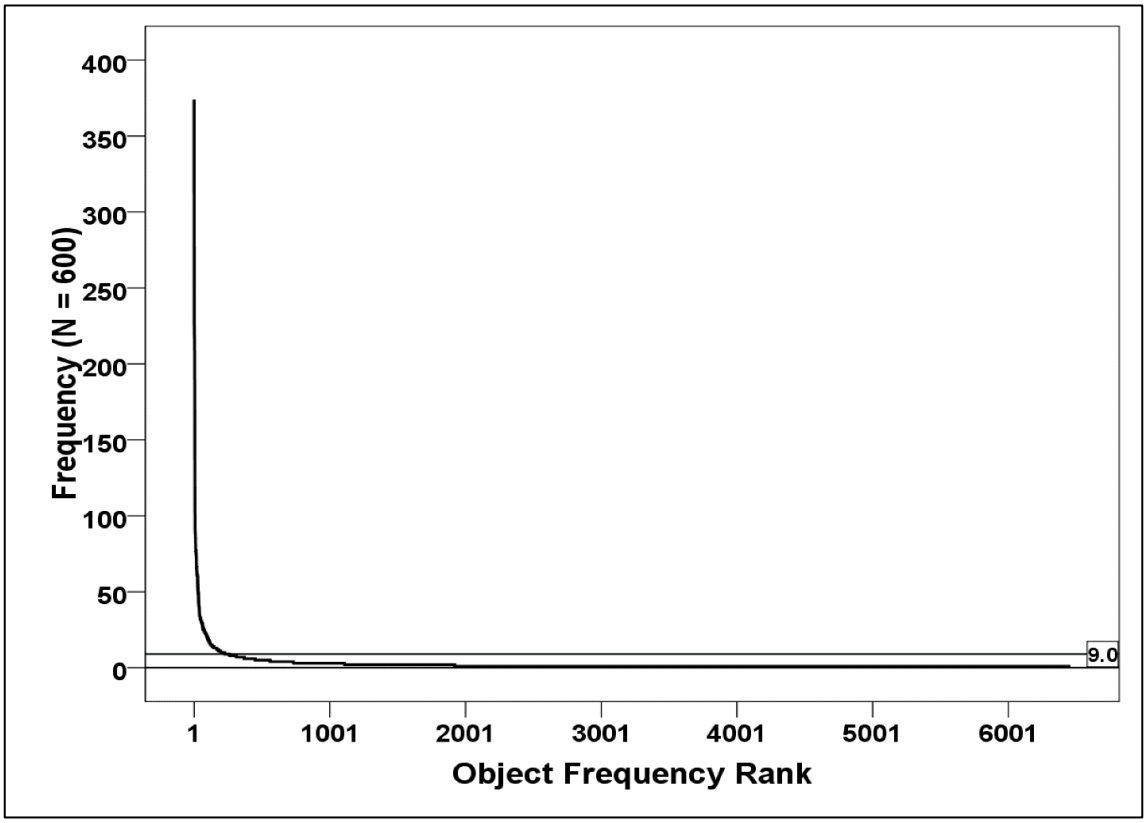

Figure 1. The frequency by rank distribution of 6,459 objects perceived in response to the Rorschach task in a sample of 600 individuals from Brazil, showing both the embedded structure of the task (the near-vertical left arm) and the long, idiographically rich set of unique responses people provide (the near-horizontal right tail).

frequently as the most common word, which in English is the word "the." The third most common word, which is "and" in English, occurs about one third as frequently as the most common word. The fourth most common word, which is "a" in English, occurs at about one quarter of the frequency as the first word, and so on. There is thus a lawful relationship between rank and frequency. If you plot that distribution of word frequency by rank order it creates a distinctive non-normal distribution; one that characterizes many natural phenomena, like the magnitude of earthquakes, the size of cities around the world, the distribution of wealth across individuals, the frequency with which particular websites are visited, the circumference of the trunk and branches of a tree, the size of craters on the moon, as well as citations to scientific papers (e.g., Clauset, Shalizi, & Newman, 2009; Newman, 2005; see also Gladwell, 2008).

These very different phenomena all create this same kind of distribution, illustrated in Figure 1, which is the distribution of Rorschach objects that were perceived in a sample of 600 individuals from Brazil (Meyer, Viglione, Mihura, Erard, & Erdberg, 2011; Villemor-Amaral, Yazigi, Nascimento, Primi, & Semer, 2007). On the vertical axis is the frequency count of objects. On the horizontal axis is the rank of those objects. There is one object that was seen about 375 times, which is at the rank of number 1 on the left side of the horizontal axis. Thus, the frequency of 375 forms the top of the vertical arm of the graph on the left hand side. A total of 6,459 different objects were reported by these 600 Brazilian non-patients. The near-vertical left arm is the embedded structure of the test. Those high frequency objects are the objects that Rorschach built into the inkblot stimuli. They are the bats, the butterflies, the people, the bears, and so on, that are structured into the inkblots with intentionally suggestive shapes. That is, the inkblots are designed to look pretty much like these things, such that they appear conventional to most people. Interestingly, there are just 30 percepts that are reported by at least 50 people in this sample. Another way to say this is that there are just 30 things that are high frequency "embedded-structure" objects in the inkblot stimuli. There is a total of 262 percepts that were reported by at least 9 people. That benchmark is important. There are 600 people in this sample, so objects seen by at least 9 people are objects reported by 1.5% of the sample. That frequency of 9 people or 1.5% of the sample is plotted as a horizontal line in Figure 1. It defines the inflection point of the plotted curve; the point where the curve transitions from the near-vertical left arm to the near-horizontal right arm. The near-horizontal right tail is the long, idiographically rich set of unique responses people provide. These are the things that are seen by relatively few people. Out of the 6,459 things that were seen in this data set, a total of 4,538 of those things were reported by just one person. Thus, 70% of the percepts that were reported were identified by just one individual, which is pretty remarkable.

When we ask people to look at the inkblots and to solve the problem of deciding what it might be, that sets in motion a range of cognitive and emotional processes where the person is engaged in trying to figure out what these stimuli might be, what they look like.

As people respond across all 10 cards it provides standardized behavioral observations. That is, answering the question using these fixed inkblot stimuli provides a sample of behavioral problem solving that addresses both the uniqueness of the individual, and their ability to perceive the somewhat nebulous external structure that is present but not fully evident in the stimulus. This problem solving process provides three elements: (1) visual attributions to the stimuli; (2) verbal and nonverbal communications about them; and (3) interactive behaviors with the examiner, with the inkblots themselves, and to some extent with the external assessment setting.

The task provides a very unique context. The only demand on the individual is to provide a verbal response to the visual stimuli. As I have said, Rorschach built into these visual stimuli considerable structure as well as contradictory "critical bits" that suggest and are evocative of many different things, but there is no thing in these inkblots that is perfectly represented.

Thus, the task engages processes that are very different than introspective self-report. The task does not require the higher-level cognitive processes that are necessary for using questionnaires. Those measures require the respondent to reflect on their personal characteristics; compare their characteristics to others in their magnitude, their frequency or generality; and then decide how much of their conclusion they really want to share with the examiner or researcher about their personal characteristics (Bornstein, 2011; Meyer, 1997).

The Rorschach on the other hand puts the responsibility for this kind of comparative classification on the examiner who has to classify the attributions, communications, and interactions along dimensions of interest (Meyer, 1997). The examiner then aggregates those coded features across all responses to compare what this individual sees, says, and does to what others see, say, and do when confronted with the same problem solving task.

Using this norm-based information one can infer likely behavior for a person in everyday life. In particular, the target behavior for these inferences concern *what the person likely does when left to his or her own predilections to understand, represent, and make meaning about complex environmental stimuli* – stimuli that validly can be seen from multiple and often very different perspectives. These environmental stimuli include internal experiences, as well as external experiences. Internal experiences encompass intrapsychic processes, like thoughts, feelings, impulses, and physiology. External experiences relate to perceptions of the external environment and they encompass social relations with friends and peers and close relations with significant others or intimate partners, as well as management of developmental demands for achievement and mastery.

So, the task allows one to see what the person does, not learn what she thinks she does. It is a reasonably brief, portable, behavioral experiment that provides "performance assessment." As a behavioral task, the most valid inferences are those in which the behaviors observed and coded in the microcosm of the task generalize to parallel mental, verbal, perceptual, and interactive behaviors in the external environment (Viglione & Rivera, 2013).

The place of Rorschach data in an assessment is rather unique. The coded behaviors can and often do reflect implicit qualities that are not recognized or reported by the respondent. In this regard the Rorschach is very much like other performance tasks of intelligence or memory. People do not really have a good sense of how good their memory is or how well they perform on memory tasks relative to others (Beaudoin & Desrichard, 2011). This is implicit information and so again it is a complement to consciously recognized self-reported characteristics. As noted above, at best Rorschach scores are just modestly correlated with parallel self-report measures (e.g., Meyer, Riethmiller, Brooks, Benoit, & Handler, 2000; Mihura et al., 2013). Consequently, valid Rorschach scores provide unique personality data and they add incrementally and meaningfully to self-report.

This largely discursive overview of the task and reasons why the Rorschach can be a useful instrument for individual differences research provides a conceptual foundation. I focus next on selected validity research that provides an empirical foundation for using the task.

An empirical foundation for Rorschach derived scales: Meta-analyses

About 15 years ago Meyer and Archer (2001) summarized the meta-analytic evidence that was available at the time on the global validity of Rorschach scores, drawing on data compiled by Atkinson (1986), Parker, Hanson, and Hunsley (1988), Garb, Florio, and Grove's (1998) reanalysis of Parker et al., and Hiller, Rosenthal, Bornstein, Berry, and Brunell-Neuleib (1999; see also Rosenthal, Hiller, Bornstein, Berry, & Brunell-Neuleib, 2001). Each of these meta-analyses compared the validity of the Rorschach to the validity of the Minnesota Multiphasic Personality Inventory (MMPI), a commonly used self-report measure of psychiatric symptomatology. Meyer and Archer also included a greatly expanded set of effect sizes from Parker et al.'s database that had never been published before.[1] At the level of hypothesized effects, they found equivalent validity, with both measures having an average r = .32 (based on 523 effect sizes for the Rorschach and 533 effects for the MMPI). Both measures also had the same average effect size when cross-method validity coefficients were aggregated within samples (r = .29, based on 73 samples for the Rorschach [N = 6,520] and 85 samples for the MMPI [N = 15,985]). These data clearly supported the general validity of the Rorschach.

However, these data were limited because they indicated that the Rorschach was *generally* valid, and generally as valid as the MMPI, but they did not indicate which *specific scores* were valid and which were not. This fact led Garb

[1] For the Rorschach, Parker et al. reported convergent validity based on 13 effect sizes from five studies (N = 283); Meyer and Archer reported convergent validity based on 247 effect sizes from 43 studies (N = 4,807). For the MMPI, Parker et al. reported convergent validity based on 66 effect sizes from 30 studies (N = 4,980); Meyer and Archer reported convergent validity based on 296 effect sizes from 58 studies (N = 11,531).

to call for a moratorium on the use of the Rorschach in clinical and forensic settings until that kind of specific meta-analytic data were in hand (Garb, 1999). Providing such extensive meta-analytic data is not an easy undertaking. However, Mihura et al. (2013) completed a review of the published literature on all the specific variables in the Rorschach Comprehensive System (Exner, 1974; 2003). The results were published in *Psychological Bulletin*, a prominent outlet for systematic reviews that had not published anything focused on the Rorschach for 25 years – since Parker et al.'s (1988) meta-analyses.

The review focused on the Comprehensive System because it was the most commonly used scoring system in the United States and in many other countries (Meyer, Hsiao, Viglione, Mihura, & Abraham, 2013). That review reliably identified all instances of validity coefficients in the published literature that had been hypothesized by any authors ($\kappa = .90$). Thus, when one author hypothesized that a particular score should be associated with a particular criterion, the review included all instances when any author had studied that same score in relation to the same criterion, regardless of their hypotheses. A total of 3,106 findings were identified at this stage. These then were reliably classified with respect to their construct relevance ($\kappa = .79$) and winnowed to identify the core findings that spoke directly to the validity of the constructs that these variables were supposed to be measuring. This reduced the number of effect sizes to a little more than 1,100. As expected, Rorschach scores were globally more strongly associated with externally-assessed criteria than with self-reported characteristics. After mapping Rorschach assessed characteristics to parallel self-reported characteristics, the average validity coefficient was .08. However, using externally assessed characteristics, like observer rating, psychiatric diagnosis, or other life-event type criteria, the average correlation between the Rorschach scales and the criterion measures was .27 on average, which is a substantial coefficient for independently assessed hetero-method validity (Hemphill, 2003; Meyer et al., 2001).

More importantly, the review was able to generate effect sizes for 53 of the 65 variables that form the interpretive core of the Comprehensive System. Relative to externally assessed criteria, 13 of the variables had excellent support, 17 had good support, 10 had modest support, 13 had no support, and the remaining 12 variables had no construct-relevant validity studies in the peer-reviewed literature. The results from these analyses also formed the core psychometric foundation for the Rorschach Performance Assessment System (R-PAS; Meyer et al., 2011), which was designed as an improvement over and replacement for the Comprehensive System.

Concurrent with the meta-analyses, Meyer et al. (2013) undertook a large scale survey with experienced clinicians who used the Rorschach. These clinicians were asked to ignore what they had been taught about Rorschach scores and to focus instead on what scores they thought were working validly or not based on their clinical practice. The survey focused heavily on Comprehensive System codes but also collected data about multiple other scoring systems. Ultimately, input was obtained from 246 clinicians working in 26 different countries. Although any single clinician did not agree strongly with any other clinician, in composite form their ratings provided very reliable judgments about what scores were and were not valid (average $\alpha = .95$). Interestingly, across the 65 variables studied by Mihura et al. (2013), the correlation between the average ratings of validity from the clinicians and the classification of the variable as having no, modest, good, or excellent validity based on the meta-analytic results was .51. This correlation indicated that the aggregated judgments of these experienced clinicians were strongly associated with the previously un-summarized findings from the research literature. However, both the aggregated clinical judgments and the meta-analytic research findings diverged from the existing authoritative review of validity for these variables, which was Exner's (2003) text on the Comprehensive System. That text essentially endorsed the validity of all Comprehensive System variables.[2] Thus, the clinicians and the meta-analytic research were diverging from Exner while converging on the same two conclusions: that some variables lacked validity and probably should not be used in clinical practice and that other variables were valid and should be emphasized in clinical practice.

In response to the Mihura et al. (2013) meta-analyses, the self-described "Rorschach critics" (see Wood et al., 2003) published a follow-up comment (Wood, Garb, Nezworski, Lilienfeld, & Duke, 2015). They made two noteworthy statements, given the years of debate associated with the Rorschach. First, they said, "The estimated validity coefficients reported by the authors [i.e., Mihura et al. (2013)] provided an unbiased and trustworthy summary of the published literature." In addition, they rescinded the global moratorium Garb (1999) had called for on use of the Rorschach in clinical and forensic settings. Wood et al. (2015) said, "He [i.e., Garb] and the other authors of this Comment agree that, in light of the compelling evidence laid out by Mihura et al., the time has come to withdraw this recommendation so far as it applies to the [14 variables that Wood et al. (2015) call the] Cognitive Quartet."

Although these statements hold positive appeal for Rorschach users, their comment also presented data concerning four variables for which they argued that if one added unpublished dissertations to the peer-reviewed literature considered by Mihura et al. (2013), observed validity coefficients might decrease or even evaporate. In response, Mihura, Meyer, Bombel, and Dumitrascu (2015) identified numerous problems with Wood et al.'s (2015) comment, including many illogical and biased arguments and many errors that they made with the data and analyses presented in their comment.

Mihura et al.'s (2013) meta-analyses only examined variables included in Exner's (2003) Comprehensive System. However, other systematic reviews and meta-analyses provide validity support for individual Rorschach variables. For the variables included in R-PAS that were not in the Comprehensive System, additional empirical support is available for the Oral Dependent Language variable (also called the Rorschach Oral Dependency scale; Bornstein, 1996; 1999), the Ego Impairment Index (Deiner, Hilsen-

[2] The two exceptions were that Exner (2003) cautioned about drawing inferences about diagnoses of depression from the Depression Index and recommended use of the Oral Dependency scale (Bornstein, 1999; Bornstein & Masling, 2005) over the CS Food score when issues of dependency were of interest.

roth, Shaffer, & Sexton, 2011), the Mutuality of Autonomy Scale (Graceffo, Mihura, & Meyer, 2014; Monroe, Diener, Fowler, Sexton, & Hilsenroth, 2013), and Space Reversal and Space Integration (Mihura, Dumitrascu, Roy, & Meyer, 2017). Somewhat paradoxically, as a function of the ongoing controversies concerning the Rorschach, the variables included in R-PAS now have more meta-analyses documenting their construct validity than the variables included in any other multiscale assessment measure, such as the MMPI or the Wechsler scales.

An empirical foundation for Rorschach derived scales: Example research

Conventionality of perception via Form Quality

The Rorschach is the only task with normed data on the conventionality of one's perceptions. Although there are numerous maximal performance neuropsychological tests that assess visual discrimination or visual-spatial processing (e.g., Lezak, Howieson, Bigler, & Tranel, 2012), the Rorschach uniquely assesses the conventionality of perception rather than visual processing skills. It does so via Form Quality variables, with the typical classifications being FQo to designate perceptions that are "ordinary" or conventional, FQu to designate perceptions that are "unusual" or atypical, and FQ– to designate perceptions that are "minus" or distorted. Many studies have documented the validity of form quality scores for assessing problems on the psychotic disorder spectrum. These variables have been unequivocally endorsed as valid by even the major critics of the Rorschach, including Dawes (1999) and Wood, Nezworski, and Garb (2003). In the Mihura et al. (2013) meta-analyses, they produced some of the strongest validity coefficients, with correlations above .45 for assessing psychotic disorder criteria.

The Rorschach Performance Assessment System developed revised tables to designate the level of form quality that should be assigned to exemplar objects at various locations within each inkblot (Meyer et al., 2011). To do so, the conventionality of perception was decomposed into its two elements, fit and frequency. Fit concerns the adequacy of the match between the object identified and the gestalt contours at the location where it was perceived. Frequency concerns how often that particular object is spontaneously mentioned at a particular location by people who are completing the task, which ties directly to the Zipf distribution in Figure 1. The R-PAS authors applied an iterative process of refinement, such that fit and frequency classifications were applied to objects in the form quality tables, responses were coded according to those criteria, and then validity was evaluated using two clinical samples and normative data.

Fit was determined by Form Accuracy ratings, in which 13,031 objects culled from a wide range of Rorschach publications in the U.S. and Europe were rated 9.9 times on average by a panel of 569 judges. These judges came from multiple countries including Brazil, China, Finland, Israel, Italy, Japan, Portugal, Romania, Taiwan, Turkey, and the United States. The judges used a 5-point scale ranging between "1. No, I cannot see it at all. Clearly it is a distortion," through "3. A little, if I work at it, I can sort of see that," to "5. Definitely … it looks exactly or almost exactly like that." The Fit ratings were then averaged for each object and each object was classified as falling at one of three levels. Objects with an average rating of 2.4 or less were considered a distortion or a poor fit to the inkblot. Objects with an average rating of 3.5 or higher were considered to be good perceptions; an object that matches that inkblot location pretty well. And then there was a mid-range between those two.

Frequency data were culled from five sets of adult Form Quality tables developed in Argentina (Lunazzi et al., 2011), Brazil (Villemor-Amaral et al., 2007), Italy (Parisi, Pes, & Cicioni, 2005), Japan (Takahashi, Takahashi, & Nishio, 2009), and Spain (Miralles Sangro, 1996; 1997). All five datasets had exact frequency counts for each object, and all objects that were reported by 1.5% or more of the people in a sample were translated into English. These were thus all the objects above the inflection point on the Zipf curve in Figure 1, which identifies the structural elements of what Rorschach built into the inkblots. Those objects were linked across the different frequency tables and then also linked to the 13,031 objects that had been rated for Form Accuracy. To use these data, the R-PAS authors emphasized international generalizability for how often the same object was seen in different samples. Like with the form accuracy ratings, three categories were formed. The least weight was assigned to objects that were not reported by 1.5% of the people in any samples, the most was given to objects that were reported by 1.5% or more of the people in at least 2 samples, and modest weight was given to those objects reported that often in just one of the samples.

To make the table more manageable for users, the authors focused on just 5,060 objects derived from the Comprehensive System Form Quality tables, which had been modified by members of John Exner's Rorschach Research Council to reduce irregularities, inconsistencies, obvious omissions, and redundancies. The final Form Quality codes were then determined by these three sources of information: the pre-existing Comprehensive System Form Quality codes, the object Frequency as seen in 5 international samples, and then the average Fit ratings from an international panel of judges. In the end, relative to the Comprehensive System classifications, about 40% of the 5,060 objects had a different classification in the new tables.

By using a large amount of internationally collected data to develop the new tables, the R-PAS developers hoped the revised classifications would more comprehensively capture contemporary and cross-culturally generalizable elements of perception. To date, two studies have examined the validity of these new tables since their publication.

Meyer and Eblin (2015) assessed 72 patients at a maximum security psychiatric inpatient facility in the U.S. These inmates were hospitalized either because they were deemed not guilty by reason of insanity or because they were deemed incompetent to stand trial because they were too psychotic. The patients were largely diagnosed with schizophrenia or schizoaffective disorder. Ratings about each patient were obtained from their treating psychiatrist and primary clinician on the Positive and Negative Syndrome Scale (Kay, Fiszbein, & Opler, 1987) following a treatment team meeting during which the patient was dis-

cussed. The ratings addressed cognitive disorganization, which encompassed conceptual disorganization or an inability to maintain focused attention; they further addressed positive symptoms of schizophrenia, which included delusions, hallucinations, grandiosity, and unusual thought content; and finally they addressed a composite measure formed from both factors. The proportion of Rorschach responses classified as having distorted form quality was strongly correlated with the disorganized factor (r = .55), moderately with the positive factor (r = .30), and strongly with the composite (r = .42). In contrast, the proportion of Rorschach responses classified as having conventional form quality had strong negative correlations with each of these criteria (disorganized, r = -.44; positive, r = -.42; composite, r = -.47). Thus, there was good evidence for validity of the new form quality tables.

Su et al. (2015) looked at the cross-cultural generalizability of R-PAS for assessing severity of disturbance in Taiwan using 90 adults who varied in their clinical status. Fifteen were non-patients, 37 were outpatients with various diagnosis, 11 were in long term day-treatment because they had more severe illnesses, and 27 were inpatients. Protocols were administered using R-PAS guidelines, but all protocols were independently coded by two different groups using either the R-PAS manual or Compressive System criteria. The criterion variables were scales from the Positive and Negative Syndrome Scale, the Magical Ideation Scale (Gross, Silvia, Barrantes-Vidal, & Kwapil, 2012), which is a measure of unusual beliefs and unusual experiences, a global clinical rating of severity, and a severity index based on the diagnosis or diagnoses assigned to patients. The variable of distorted form quality coded using the R-PAS form quality tables had an average correlation with these criteria of .48, while the corresponding variable based on the Comprehensive System form quality tables had an average correlation of .39. In regression analysis, the R-PAS tables had incremental validity over the Comprehensive System tables for all criteria, but the reverse was never true. Thus, the R-PAS Form Quality tables worked better; they provided valid information for coding perceptual accuracy in Taiwan that could not be obtained from the Comprehensive System Form Quality tables.

Unique vs. common responses: MRI and fMRI findings

Shifting focus, Asari and colleagues (2008; 2010a; 2010b) conducted three interrelated studies examining unique perceptions in the Rorschach using functional magnetic resonance imaging (fMRI) or structural magnetic resonance imaging (MRI). Unique perceptions were defined as objects that were reported by people in their study sample of 68 students but not reported in a normative control sample. Thus, unique perceptions would be coded as having either unusual or distorted form quality and they would fall on the long horizontal tail of the Zipf distribution in Figure 1.

In 2008, Asari et al. showed that common or frequent perceptions were associated with the left anterior frontal cortex and the visual cortex in the occipital lobes, while unique perceptions were associated with the right temporal pole. Given the links between the right temporal pole and limbic structures, these findings suggested that unique perceptions were associated with emotional reactions, perhaps particularly of an autobiographical nature. The findings provided some neuroimaging foundation for the clinical belief that unusual perceptions reflect instances when personal reactions or personal conflicts override the ability to perceive experiences in a conventional manner.

Asari et al. (2010a) then found that people who gave more unique perceptions had an enlarged amygdala and cingulate gyrus, both of which are part of the limbic system. Enlarged areas of the brain suggest that they are being used more frequently, so the findings again suggested that personal, emotional reactions contributed to unique perceptions. Finally, Asari et al. (2010b) attempted to integrate results from the two earlier studies. They used functional connectivity analysis of fMRI data to determine the role of the amygdala in generating unique perceptions. The amygdala generated a positive, excitatory link between the right temporopolar region previously mentioned and the left anterior prefrontal cortex. Thus, people with an active amygdala had a link between that temporal pole and the anterior prefrontal cortex. This suggested that personal, autobiographically relevant emotional reactions helped generate the thoughts or ideas about what was in the inkblot that was classified as a unique perception. At the same time that the amygdala was doing this, it was generating negative inhibitory connections from the right temporopolar region to the occipital regions. As such, personal and autobiographically relevant emotional reactions helped to simultaneously inhibit general visual processing. That is, it impaired general perception of the environment. Thus, amygdala activation impaired the typical perception of the external environment that is processed in the occipital lobes, while it facilitated the perception of personalized, unique images in the inkblots.

The data show how personally relevant, unusual representations take precedence over standard visual imagery. In essence, the personally relevant, unique perceptions "force themselves" into an inkblot representation, rather than allowing the person to take in the actual visual cues that are present in the inkblot stimuli. Together, these studies provided important data for understanding the response process foundation for Rorschach assessment of perceptual conventionality and idiosyncrasy.

The mirror neuron system and human movement responses

Giromini and colleagues have completed a sequence of studies looking at the link between perceptions of human activity and movement on the Rorschach and the mirror neuron system (Ando' et al., 2015; Giromini, Porcelli, Viglione, Parolin, & Pineda, 2010; Pineda, Giromini, Porcelli, Parolin, & Viglione, 2011; Porcelli, Giromini, Parolin, Pineda, & Viglione, 2013; see also Porcelli & Kleiger, 2016). The mirror neuron system represents the cortical firing that takes place in primates either when an action is executed or when an action is observed taking place. In humans, the mirror neuron system is thought to reflect the neurophysiological manifestation of an action. That is, it is an internal psycho-motor representation of physical activity taking place or being observed. The mirror neuron system also reacts to and represents the underlying intentions,

thoughts, and feelings that motivate an action. Thus, when people observe human action taking place they have a mental template of what intentions, thoughts, and feelings motivate those behaviors in the actor. This suggests a potential neurological basis for empathy, for theory of mind – that is for understanding how other people think about or process their experiences, and for the ability to process facial emotion.

The first study in this series used a small sample of 19 undergraduates (Giromini et al., 2010). The authors recorded activity in the human mirror neuron system by measuring activity in the mu frequency band of an electroencephalograph (EEG) recording of cortical activity taking place over the sensorimotor cortex. When at rest, the sensorimotor neurons fire in synchrony. This leads to large amplitude EEG oscillations in the range of 8 to 13 hertz, which is the mu frequency band. When performing an action, those same neurons fire asynchronously and thereby decrease the power of the mu-band EEG oscillations, causing cortical firing to fall from the 8-13 hertz range to a lower frequency. Giromini et al. collected EEG data at baseline and then during three other conditions. At baseline they had people look at the middle of a white card. In the first condition, which followed standard assessment procedures, participants spontaneously gave responses to four Rorschach inkblots where perceptions of human movement are either common or uncommon, and attributions of human activity were contrasted with attributions that did not contain human activity. In the second condition, participants were given the same inkblots but asked to identify commonly reported human movement perceptions or commonly seen static objects, depending on the card. In the third condition, participants viewed representational drawings of the inkblots, two with humans engaged in apparent movement, and two with static objects. In each condition, Giromini et al. found greater mu suppression when human movement was present than when it was absent. For the standard assessment condition, the findings produced a large effect size (r = .51). Thus, the data suggested that giving a response with human movement, identifying suggested human activity, or perceiving stimuli depicting human movement was associated with activity in the mirror neuron system.

A second study had 24 undergraduates view all 10 inkblots using standard administration (Pineda et al., 2011). The researchers compared spontaneously given human movement responses to non-human movement responses using the same EEG procedures as before. They found that when people spontaneously give human movement responses relative to other kinds of responses, there is a suppression in mu, indicating an increase in motor neuron activity, this time producing a very large effect size of r = .67. Porcelli et al. (2013) further delineated the findings in this sample, showing that mu suppression was present for human movement but not for human content without movement, movement by animals or inanimate objects, or use of the inkblot color, shading, or form. Thus, generating Rorschach responses with human movement appeared to distinctively activate the mirror neuron system.

In the next study Ando' et al. (2015) had 36 undergraduates undergo inhibitory repetitive transcranial magnetic stimulation (rTMS), a procedure that temporarily interferes with neuronal activity. Prior work with rTMS showed that interfering with the left inferior frontal gyrus, or the motor strip, decreased performance in empathy related tasks and it eliminated the mu suppression of the EEG signal that is typically observed on empathy related tasks. Ando' et al. hypothesized that this experimental procedure would suppress human movement on the Rorschach as well. They conducted a repeated measures study with people randomly assigned to one of two conditions. In the control condition participants were tested at baseline using standard administration and then retested with rTMS at the top of the head (the vertex), and not over the motor strip. In the experimental condition participants received standard administration at baseline and at retest the rTMS was over the left inferior frontal gyrus. Both groups were assessed with three Rorschach cards that pulled for frequent human movement. Ando' et al. found very strong and clear effects. The control group did not change in their frequency of human movement responses from baseline to the stimulation condition. The experimental group had a typical number of human movement responses at baseline, but those responses were markedly suppressed when their motor strip was temporarily blocked by magnetic interference, which made it difficult for them to empathize with representations of human activity. This led to a very large effect size, with the relevant means being almost two standard deviations apart (d = 1.96).

In summary, these three small studies using two very different methods suggest that the process of seeing human movement is associated with mirror neuron activation. When someone produces a human movement response, they are implicitly identifying with and experiencing the action that is taking place. This supports the historical belief that human movement responses are related to empathy, thoughtful reflection, and introversiveness. These responses also are associated with theory of mind, or the capacity to put oneself in somebody else's shoes and understand what it is they are experiencing. More recently, Giromini et al. (2017) have replicated the same findings in a small MRI study.

Tactile imagery, attachment, and texture responses

In a series of studies, Iwasa and Ogawa (2010; 2013; 2016) examined Rorschach responses in which variation in the saturation of ink contributed to a tactile attribution, which is known as a shading-based Texture response. They attempted to both find the correlates of the Texture response with self-reported attachment styles and to articulate the response process foundation of tactile representations. In an initial sample of 47 students, Iwasa and Ogawa (2010) found people who had one Texture response felt more securely attached than people who had more than one Texture response or no Texture in their protocol. People who had more than one Texture response in their protocol reported being more preoccupied with attachment security, having more attachment anxiety, and having less attachment security than the people who had one Texture response. Unexpectedly, the people with no Texture in their protocol reported being like the people with more than one response; they were more preoccupied, had more attachment anxiety, and felt less security than the people with

one Texture response.

Iwasa and Ogawa (2013) reported 2 studies. The first was a sample of 20 people who were administered the Rorschach in addition to a scale assessing the vividness of tactile imagery. They anticipated that the number of Texture responses would be associated with reports of more vivid tactile images, and they found a strong effect supporting their hypothesis (r = .52). The second study examined self-reports of the vividness of tactile imagery, strength of emotional reactions to tactile imagery, and attachment styles. The authors found that tactile vividness was predicted by the strength of emotional reactions to tactile imagery. Further, the strength of emotional reactions to tactile imagery mediated the emotion suppressing effects of attachment avoidance. People with avoidant attachment styles have more muted emotional reactions to tactile imagery, and this is carried forward as less vividness in mentally picturing tactile experiences. These self-report findings suggested that attachment avoidance may also inhibit Texture responses on the Rorschach.

Finally, Iwasa and Ogawa (2016) used a novel task that allowed them to assess the extent to which people responded to priming by tactile visual images. Participants (N = 35) were shown tactile visual images and the researchers quantified the extent to which those images facilitated their ability to classify words. That is, they assessed the accessibility of tactile imagery in one's thinking. Importantly, this was not consciously reported accessibility but data based on actual performance. They found that accessibility of tactile knowledge, attachment anxiety, and Texture responses were all strongly correlated (rs $\geq$ .50). In addition, the propensity to have accessible tactile imagery partially mediated or carried forward the effects of attachment anxiety on Texture responses, such that more anxiously attached people were more prone to have accessible tactile knowledge, and in turn this more accessible tactile knowledge led to more Texture responses on the Rorschach. In this study, although attachment avoidance was expected to inhibit tactile awareness and texture responding, it showed no significant effects.

Establishing normative benchmarks: What people typically see, say, and do

In 2007 a special supplement to the *Journal of Personality Assessment* was published containing internationally collected normative data for the Comprehensive System (Shaffer, Erdberg, & Meyer, 2007). The supplement contained 39 samples of adults and children or adolescents, about evenly split, drawn from many different countries, including Argentina, Australia, Belgium, Brazil, Denmark, Finland, Greece, Israel, Italy, Japan, The Netherlands, Peru, Portugal, Romania, Spain, and the U.S. In an article summarizing these findings, as well as through published adult data from France, Meyer, Erdberg, and Shaffer (2007) showed how adults around the world looked very similar to each other, such that pooling the results to form composite international reference values was justified. Thus, because scores generalized across languages, cultures, and strategies for recruiting non-patients, these norms provided benchmark standards for what people typically see, say, and do when completing the Rorschach task. Knowing that the Rorschach produces generalizable norms like this makes the Rorschach rather unique and distinctive among psychological assessment measures.

However, the data also demonstrated that Exner's (2003) standard normative reference data for the Comprehensive System were incorrect for some key variables. The internationally collected data both differed from and corrected the standard Comprehensive System norms. Early evidence suggested that Exner's norms might be problematic (Shaffer, Erdberg, & Haroian, 1999; Viglione & Hilsenroth, 2001; Wood, Nezworski, Garb, & Lilienfeld, 2001), although in many ways it appeared that the problems were limited to Form Quality variables (Meyer, 2001a). However, with the synthesized international data it became evident that Exner's standard norms erroneously caused non-patients to appear psychologically unhealthy on about one third of the variables (Giromini, Viglione, & McCullaugh, 2015; Meyer, Shaffer, Erdberg, & Horn, 2015; Meyer, Viglione, & Mihura, 2017; Viglione & Giromini, 2016). Although criticisms can be and have been raised about the quality of all the studies that contributed to the international norms, even when those norms were limited to data collection efforts that relied on highly trained examiners in carefully designed and executed studies, the same differences from Exner's standard were evident (Meyer et al., 2015). Thus, the composite of data indicate that clinicians using the standard Comprehensive System norms will incorrectly infer that non-patients are prone to perceptual distortions, see the world in an atypical idiosyncratic manner, tend to be simplistic, lack affective resources, lack coping resources in general, are prone to affective disturbances and dysregulation, and misunderstand others and misperceive relationships. The internationally collected norms correct for those problems and those norms should therefore be used in clinical practice.

Ego depletion and complexity variables

In closing, the next three sections review recent and very different individual studies that speak to the utility of the Rorschach for personality research. The first examined ego depletion, which is believed to occur because people have a finite pool of cognitive resources for engaging in complex mental processing (Baumeister, Bratslavsky, Muraven, & Tice, 1998). Meta-analytic data have supported the ego depletion model, indicating that acts of effortful self-regulation temporarily deplete that pool of resources, such that on a subsequent self-regulation task, cognitive functioning is impaired (Hagger, Wood, Stiff, & Chatzisarantis, 2010; though see Hagger et al., 2015; Tuk, Zhang, & Sweldens, 2015).

Charek, Meyer, and Mihura (2016) compared a control group of 43 students who completed a simple letter cancellation task to an experimental group of 54 students who completed a complex letter cancellation task designed to deplete their cognitive reserves. The authors anticipated that the Rorschach protocols for the depleted students would show more reactivity to color, less cognitive sophistication, and more frequent lapses in how they logically put visual images together, in addition to self-reporting greater fatigue and less attentiveness. Their hypotheses were partially supported. Despite a surprising absence of self-

reported differences, ego-depleted participants had Rorschach protocols with lower scores on two variables indicative of sophisticated combinatory thinking (responses that encompassed the whole inkblot while simultaneously making meaningful connections between distinct objects, and instances when the white background of the card was integrated with the inkblot proper). They also had higher levels of color receptivity. The strongest effect was that the depleted participants had lower scores on a composite variable computed across all hypothesized markers of cognitive complexity. Depletion did not affect three R-PAS variables that were hypothesized to be unrelated to cognitive complexity. In addition, baseline levels of self-reported achievement striving moderated the effect of the experimental manipulation on color receptivity. Participants who described themselves as high in achievement striving threw themselves more fully into the depleting tasks, showing higher levels of attentiveness. Subsequently, when completing the Rorschach they were the ones most likely to mentally coast and provide less global synthetic processing and more color reactivity. This produced a clear interaction effect of achievement striving and conditions on color reactivity. Those people with the highest levels of achievement striving were the least reactive to color in the control condition. However, after putting greater effort into completing the cognitively depleting task, they were the most reactive to color in the depletion condition. Overall, this study, though limited by a less potent experimental intervention than planned, provided clear support for color reactivity, synthetic cognitive operations, and the overall degree of complexity that can be found in a protocol.

Cardiovascular risk and aggressive imagery

Meyer, Katko, Mihura, Klag, and Meoni (2017) examined hostility as a predictor of premature cardiovascular disease and coronary heart disease in 416 men from the Johns Hopkins Precursors Study. As medical students attending Johns Hopkins University from 1948 to 1964, they completed the Rorschach and a self-report inventory created specifically for this study in the 1940s. They were then followed in medical school and every year since graduation up to the present time. The sample consisted of all the men who had been individually administered the Rorschach at baseline and produced at least 18 responses. The Rorschach protocols were coded for Aggressive Content, which encompasses imagery of powerful, threatening, dangerous, or predatory creatures or objects. The self-report inventory was scored for three items dealing with anger in response to stress. Although it was initially predicted that low self-reported anger in combination with high aggressive imagery on the Rorschach would produce the greatest risk for early cardiovascular disease and coronary heart disease over the subsequent 30-year period, the combination of high self-reported anger and high aggressive imagery conferred the greatest risk for these disease endpoints. The psychological variables also incrementally predicted these outcomes over baseline covariates that are considered medically important to the development of cardiovascular disease, including smoking, serum cholesterol, systolic and diastolic blood pressure (both at rest and after stress), heart rate (both at rest and after stress), body mass index; alcohol use, whether the student's mother or father had developed premature cardiovascular disease, and the occupation of the student's mother and father. The only covariates that also significantly contributed to disease prediction were smoking for cardiovascular disease and coronary heart disease, and serum cholesterol for coronary heart disease. Thus, men who experienced angry reactions to stress and who had dangerous, threatening, aggressive imagery on their mind while in medical school, subsequently developed cardiovascular disease and coronary heart disease more quickly than other men, and this predictive capacity added novel information about risk over a range of empirically important medical predictors.

Manifestations of grandiosity and narcissism

Finally, another recently published article encompassing two studies attempted to improve the assessment of grandiosity and narcissism using the imagery and behaviors expressed while completing the Rorschach (Gritti, Marino, Lang, & Meyer, 2017). This article examined seven variables previously described in the literature: Omnipotence, where the person behaves in a way to claim unrealistic power, influence, or worth; Idealization, where the person perceives grand, important, or powerful objects; Reflection, where the person sees an object and its mirrored or reflected image; Personal Knowledge Justification, where the person relies on private knowledge or experience to justify a response; Exhibitionism, where the person identifies objects engaged in activities that are performed for an audience or designed for public display; Magic, where the person perceives magical figures or objects associated with magical powers; and Elevated Mood States, where the person identifies positive affective states like fun, pleasure, or happiness either in the percept being seen in the inkblots or in himself or herself.

In addition, four new variables were created in an attempt to address other manifestations of grandiosity and narcissism (Meyer, Gritti, & Marino, 2017). Based on the idea that narcissistic and grandiose people think everything is about them, including their inkblot percepts, Expanded Personal Reference is coded when a person places themselves into a response or links themselves to their response in some way. Narcissistic Devaluation is coded when positively invested or embellished objects are also devalued, dismissed, or denigrated; Narcissistic Deflation is coded for responses where objects are missing a key part of their identity, possess deflated or impotent parts, or are described as dying, decaying, deteriorating or eroding. Finally, Narcissistic Denial is coded when the respondent denies or minimizes the impact of perceptions connected to weakness, vulnerability, inferiority, or other negative affective states.

The two studies described by Gritti et al. (2017) encompass clinical and non-clinical adult samples. The non-clinical sample consisted of 145 English language protocols from the norms used by the Rorschach Performance Assessment System (Meyer et al., 2011). The clinical sample consisted of 100 Italian outpatients. In addition to completing the Rorschach, 55 patients in the clinical sample were rated using the Shedler-Westen Assessment Procedure (SWAP-200), from which its two scales of narcis-

sistic personality disorder were obtained. Clinicians who were treating these patients rated them on the SWAP-200 after seeing the person in treatment for at least five sessions. Fifty of these patients also provided self-reported personality disorder symptoms using the Millon Clinical Multiaxial Inventory (MCMI-III), from which its narcissistic personality disorder scale was used to test incremental validity.

Analyses included principal components analysis in both studies to evaluate if the 11 variables were measuring a common dimension of narcissistic grandiosity. In the clinical sample, Rorschach scores were correlated with the SWAP-200 ratings of narcissism, and regression analyses tested the incremental contribution of the Rorschach over the MCMI. Across samples there was variability in the components that were identified. However, in both datasets there was one clear and strong component, and it was defined by the same variables in each dataset. This component was most strongly defined by four variables: Expanded Personal Reference, Idealization, Personal Knowledge Justification, and Omnipotence.

In the clinical sample that latter component correlated strongly with the clinician ratings of narcissism (average *r* = .41). The individual variables of Personal Knowledge Justification, Expanded Personal Reference, and Omnipotence also correlated with the criteria (*r*s from .29 to .51). Although Idealization defined the common Rorschach dimension, it did not correlate with clinician ratings of narcissism, probably because idealization is more of an internal process tied to attributions and imagery and it is not as socially evident as Personal Knowledge Justification, Expanded Personal Reference, and Omnipotence, all of which are interactive behaviors that are coded during the Rorschach task. The other seven variables did not correlate with the clinician ratings of narcissism. In particular, although the Reflection variable has historically been interpreted as a measure of narcissism in both the Comprehensive System and the Rorschach Performance Assessment System, it showed no relationship with the clinician ratings.

The primary Rorschach component defined by Expanded Personal Reference, Idealization, Personal Knowledge Justification, and Omnipotence also clearly incremented over the self-report scale in predicting the clinician ratings of narcissistic personality disorder. When considered together, self-reports of narcissism were uncorrelated with the clinician ratings (β = .04), but the Rorschach component had a strong association (β = .41).

Conclusion

In summary, in this review I have tried to provide a contemporary overview of Rorschach's inkblot task, including how it was developed, why it seems to work as it does, and how to contextualize inferences from the things that people see, say, and do while completing the task. I then provided a review of the meta-analyses conducted on the Rorschach, and a selection of recent research studies with the aim of showing its value as a behavioral performance task that provides a valuable complement to self-reported characteristics – both in clinical practice and in research on personality processes. Although the task is much more time consuming to use than a self-report inventory, Rorschach performance provides a very unique source of information about people and I hope to have made a reasonable case for explaining what it can add to assessing and understanding personality.

AUTHOR NOTES

This article is based on an invited address given at the 2nd World Conference on Personality, Búzios, RJ, Brasil.

The author receives royalties on the sale of the manual for the Rorschach Performance Assessment System (R-PAS) and associated products.

ACKNOWLEDGEMENT

I am grateful to Boele De Raad and Claudio Hutz for inviting this talk and to Boele De Raad for his patiently supportive encouragement to convert it into an article.

REFERENCES

Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological Bulletin, 131*, 361-382. doi: 10.1037/0033-2909.131.3.361

Ando', A., Salatino, A., Giromini, L., Ricci, R., Pignolo, C., Cristofanelli, S., Ferro, L., Viglione, D. J., & Zennaro, A. (2015). Embodied simulation and ambiguous stimuli: The role of the mirror neuron system. *Brain Research, 1629*, 135-142. doi: 10.1016/j.brainres.2015.10.025

Asari, T., Konishi, S., Jimura, K., Chikazoe, J., Nakamura, N., & Miyashita, Y. (2008). Right temporopolar activation associated with unique perception. *Neuroimage, 41*, 145. doi:10.1016/j.neuroimage.2008.01.059

Asari, T., Konishi, S., Jimura, K., Chikazoe, J., Nakamura, N., & Miyashita, Y. (2010a). Amygdalar enlargement associated with unique perception. *Cortex, 46*, 94-99. doi:10.1016/j.cortex.2008.08.001

Asari, T., Konishi, S., Jimura, K., Chikazoe, J., Nakamura, N., & Miyashita, Y. (2010b). Amygdalar modulation of frontotemporal connectivity during the inkblot test. *Psychiatry Research: Neuroimaging, 182*, 103-110. doi:10.1016/j.pscychresns.2010.01.002

Atkinson, L. (1986). The comparative validities of the Rorschach and MMPI: A meta-analysis. *Canadian Psychology, 27*, 238-247. doi: 10.1037/h0084337

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*, 1252-1265. doi: 10.1037/0022-3514.74.5.1252.

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science, 2*, 396-403. doi: 10.1111/j.1745-6916.2007.00051.x

Beaudoin, M., & Desrichard, O. (2011). Are memory self-efficacy and memory performance related? A meta-analysis. *Psychological Bulletin, 137*, 211-241. doi: 10.1037/a0022106

Bornstein, R. F. (1996). Construct validity of the Rorschach Oral Dependency Scale: 1967-1995. *Psychological Assessment, 8*, 200–505. doi:10.1037/1040-3590.8.2.200

Bornstein, R. F. (1999). Criterion validity of objective and projective dependency tests: A meta-analytic assessment of

behavioral prediction. *Psychological Assessment, 11*, 48-57. doi:10.1037/1040-3590.11.1.48

Bornstein, R. F. (2011). Toward a process-focused model of test score validity: Improving psychological assessment in science and practice. *Psychological Assessment, 23*, 532-544. doi: 10.1037/a0022402

Bornstein, R. F., & Masling, J. M. (Eds.). (2005). *Scoring the Rorschach: Seven validated systems.* Mahwah, NJ: Lawrence Erlbaum Associates.

Buchanan, T. (2016). Self-report measures of executive function problems correlate with personality, not performance-based executive function measures, in nonclinical samples. *Psychological Assessment, 28,* 372-385. doi: 10.1037/pas0000192

Charek, D. B., Meyer, G. J., & Mihura, J. L. (2016). The impact of an ego depletion manipulation on performance-based and self-report assessment measures. *Assessment, 23*, 637–649. doi:10.1177/1073191115586580

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review, 51*, 661-703. https://doi.org/10.1137/070710111

Connelly, B. S., & Ones, D. S. (2010). Another perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136*, 1092-1122. doi:10.1037/a0021212

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). New York: Harper Collins Publishers.

Cyders, M. A., & Coskunpinar, A. (2011). Measurement of constructs using self-report and behavioral lab tasks: Is there overlap in nomothetic span and construct representation for impulsivity? *Clinical Psychology Review, 31*, 965-982. doi:10.1016/j.cpr.2011.06.001

Dawes, R. M. (1999). Two methods for studying the incremental validity of a Rorschach variable. *Psychological Assessment, 11*, 297-302. doi: 10.1037/1040-3590.11.3.297

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin, 141*, 858-900. doi: 10.1037/a0038498

Diener, M. J., Hilsenroth, M. J., Shaffer, S. A., & Sexton, J. E. (2011). A meta-analysis of the relationship between the Rorschach Ego Impairment Endex (EII) and psychiatric severity. *Clinical Psychology & Psychotherapy, 18*, 464-485. doi: 10.1002/cpp.725

Exner, J. E. (1974). *The Rorschach: A Comprehensive System, Vol. 1: Basic foundations*. New York: Wiley.

Exner, J. E. (2003). *The Rorschach: A comprehensive system. Volume 1. Basic foundations and principles of interpretation* (4th ed.). New York: Wiley.

Exner, J. E., Jr. (1996). A comment on "The comprehensive system for the Rorschach: A critical examination." *Psychological Science, 7*, 11-13. doi: 10.1111/j.1467-9280.1996.tb00659.x

Garb, H. N. (1999). Call for a moratorium on the use of the Rorschach Inkblot Test in clinical and forensic settings. *Assessment, 6*, 313-317. doi: 10.1177/107319119900600402

Garb, H. N., Florio, C. M., & Grove, W. M. (1998). The validity of the Rorschach and the Minnesota Multiphasic Personality Inventory: Results from meta-analyses. *Psychological Science, 9*, 402-404. doi 10.1111/1467-9280.00075

Giromini, L., Viglione, D. J., Pineda, J., Porcelli, P., Hubbard, D., Zennaro, A., & Cauda F. (2017). *Human Movement responses to the Rorschach and mirroring activity: An fMRI study*. Manuscript submitted for publication.

Giromini, L., Porcelli, P., Viglione, D. J., Parolin, L., & Pineda, J. A. (2010). The feeling of movement: EEG evidence for mirroring activity during the observations of static, ambiguous stimuli in the Rorschach cards. *Biological Psychology, 85*, 233-241. doi:10.1016/j.biopsycho.2010.07.008

Giromini, L., Viglione, D. J., & McCullaugh, J. (2015). Introducing a Bayesian approach to determining degree of fit with existing Rorschach norms. *Journal of Personality Assessment, 97*, 354-363. doi: 10.1080/00223891.2014.959127

Gladwell, M. (2008). *Outliers: The story of success*. New York, NY: Little, Brown and Company.

Graceffo, R. A., Mihura, J. L, & Meyer, G. J. (2014). A meta-analysis of an implicit measure of personality functioning: The Mutuality of Autonomy Scale. *Journal of Personality Assessment, 96*, 581-595. doi: 10.1080/00223891.2014.919299

Gritti, E. S., Marino, D. P., Lang, M., & Meyer, G. J. (2017). Assessing narcissism using Rorschach-based imagery and behavior validated by clinician-reports: Studies with clinical and nonclinical adults. *Assessment*. Advance online publication. doi: 10.1177/1073191117715728

Gross, G. M., Silvia, P. J., Barrantes-Vidal, N., & Kwapil, T. R. (2012). Psychometric properties and validity of the short forms of the Wisconsin Schizotypy Scales in two large samples. *Schizophrenia Research, 134*, 267-272. doi: 10.1016/j.schres.2011.11.032

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A., … Zwienenberg, M. (2015). A multi-lab pre-registered replication of the ego depletion effect. *Perspectives on Psychological Science*, Advanced online publication.

Hagger, M. S., Wood, C., Stiff, C., Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin, 136*, 495-525. doi: 10.1037/a0019486.

Hardt, J., & Rutter, M. (2004). Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *Journal of Child Psychology and Psychiatry, 45*, 260-273. doi: 10.1111/j.1469-7610.2004.00218.x

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist, 58*, 78-79. doi: 10.1037/0003-066X.58.1.78

Henry, B., Moffitt, T. E., Caspi, A., Langley, J., & Silva, P. A. (1994). On the "remembrance of things past": A longitudinal evaluation of the retrospective method. *Psychological Assessment, 6*, 92-101. doi: 10.1037/1040-3590.6.2.92

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. doi: 10.1037/1040-3590.11.3.278

Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment, 15*, 446-455. doi: 10.1037/1040-3590.15.4.446

Iwasa, K., & Ogawa, T. (2010). The relationship between texture responses on the Rorschach and adult attachment. *Rorschachiana, 31*, 4-21. doi: 10.1027/1192-5604/a000002

Iwasa, K., & Ogawa, T. (2013). Rorschach texture responses are related to adult attachment via tactile imagery and emotion. *Rorschachiana, 34*, 115-136. doi: 10.1027/1192-5604/a000045

Iwasa, K., & Ogawa, T. (2016). Psychological basis of the relationship between the Rorschach Texture response and adult attachment: The mediational role of the accessibility of tactile knowledge. *Journal of Personality Assessment, 98*, 238-246. doi: 10.1080/00223891.2015.1099540

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin, 13*, 261-276. doi: 10.1093/schbul/13.2.261

Klonsky, E. D., Oltmanns, T. F., & Turkheimer, E. (2002). Informant-reports of personality disorder: Relation to self-reports and future research directions. *Clinical Psychology: Science and Practice, 9*, 300-311. doi:10.1093/clipsy.9.3.300

Lezak, M., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment (5th ed.)*. New York, NY, US: Oxford University Press.

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. doi: 10.1111/1529-1006.002

Lunazzi, H. A., Urrutia, M. I., García de la Fuente, M., Elías, D., Fernández, F., De La Fuente, S., Bianco, A. S., & Sarachu, A. (2011). *Is Form Quality (FQ) a cultural context related variable? Presentation of the Argentinean FQ Tables*. Unpublished manuscript.

Meyer, G. J. (1997). On the integration of personality assessment methods: The Rorschach and MMPI. *Journal of Personality Assessment, 68*, 297-330. doi: 10.1207/s15327752jpa6802_5.

Meyer, G. J. (1999). Introduction to the first Special Section in the Special Series on the utility of the Rorschach for clinical assessment. *Psychological Assessment, 11*, 235-239. doi: 10.1037/1040-3590.11.3.235

Meyer, G. J. (2001a). Evidence to correct misperceptions about Rorschach norms. *Clinical Psychology: Science and Practice, 8*, 389-396. doi: 10.1093/clipsy.8.3.389

Meyer, G. J. (2001b). Introduction to the final Special Section in the Special Series on the utility of the Rorschach for clinical assessment. *Psychological Assessment, 13*, 419-422. doi: 10.1037/1040-3590.13.4.419

Meyer, G. J. (2002). Implications of information-gathering methods for a refined taxonomy of psychopathology. In L. E. Beutler & M. Malik (Eds.), *Rethinking the DSM: Psychological perspectives* (pp. 69-105). Washington, DC: American Psychological Association. doi: 10.1037/10456-003

Meyer, G. J., & Archer, R. P. (2001). The hard science of Rorschach research: What do we know and where do we go? *Psychological Assessment, 13*, 486-502. doi: 10.1037/1040-3590.13.4.486

Meyer, G. J. & Eblin, J. J. (2015, February). *Initial findings from the Center for Forensic Psychiatry and University of Toledo collaboration: Assessing problems in thinking and perception*. Presented at the Center for Forensic Psychiatry, Saline, MI; February 27.

Meyer, G. J., Erdberg, P., & Shaffer, T. W. (2007). Towards international normative reference data for the Comprehensive System. *Journal of Personality Assessment, 89*, S201-S216. doi: 10.1080/00223890701629342

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., . . . Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist, 56*, 128-165. doi:10.1037//OOO3-O66X.56.2.128

Meyer, G. J., Gritti, E. S., & Marino, D. P. (2017). *Coding criteria for potential Rorschach grandiosity and narcissism variables*. Unpublished manual, Department of Psychology, University of Toledo, Toledo, OH.

Meyer, G. J., Hsiao, W., Viglione, D. J., Mihura, J. L., & Abraham, L. M. (2013). Rorschach scores in applied clinical practice: A survey of perceived validity by experienced clinicians. *Journal of Personality Assessment, 95*, 351-365. doi: 10.1080/00223891.2013.770399

Meyer, G. J., Katko, N. J., Mihura, J. L., Klag, M. J., & Meoni, L. (2017). The incremental validity of self-report and performance based methods for assessing hostility to predict cardiovascular disease in physicians. *Journal of Personality Assessment*. Advance online publication. doi: 10.1080/00223891.2017.1306780

Meyer, G. J., Riethmiller, R. J., Brooks, R. D., Benoit, W. A., & Handler, L. (2000). A replication of Rorschach and MMPI-2 convergent validity. *Journal of Personality Assessment, 74*, 175-215. doi: 10.1207/S15327752JPA7402_3

Meyer, G. J., Shaffer, T. W., Erdberg, P., & Horn, S. L. (2015). Addressing issues in the development and use of the Composite International Reference Values as Rorschach norms for adults. *Journal of Personality Assessment, 97*, 330-347. doi: 10.1080/00223891.2014.961603

Meyer, G. J., Viglione, D. J., & Mihura, J. L. (2017). Psychometric foundations of the Rorschach Performance Assessment System (R-PAS). In R. Erard & B. Evans (Eds.), *The Rorschach in multimethod forensic practice* (pp. 23-91). New York, NY: Routledge.

Meyer, G. J., Viglione, D. J., Mihura, J. L., Erard, R. E., & Erdberg, P. (2011). *Rorschach Performance Assessment System: Administration, Coding, Interpretation, and Technical Manual*. Toledo, OH: Rorschach Performance Assessment System.

Mihura, J. L., Dumitrascu, N., Roy, M., & Meyer, G. J. (2017). The centrality of the response process in construct validity: An illustration via the Rorschach Space response. *Journal of Personality Assessment.* Advance online publication. doi: 10.1080/00223891.2017.1306781

Mihura, J. L., Meyer, G. J., Bombel, G., & Dumitrascu, N. (2015). Standards, accuracy, and questions of bias in Rorschach meta-analyses: Reply to Wood, Garb, Nezworski, Lilienfeld, and Duke (2015). *Psychological Bulletin, 141*, 250-260. doi: 10.1037/a0038445

Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2013). The validity of individual Rorschach variables: Systematic reviews and meta-analyses of the Comprehensive System. *Psychological Bulletin, 139*, 548-605. doi: 10.1037/a0029406

Miralles Sangro, F. (1996). *Rorschach: Tablas de localización y calidad formal en una muestra española de 470 sujetos [Rorschach: Tables of location and form quality in a Spanish sample of 470 subjects]*. Madrid: Universidad Pontifícia Comillas.

Miralles Sangro, F. (1997). Location tables, Form Quality, and Popular responses in a Spanish sample of 470 subjects. *Rorschachiana, 22*, 38-66.

Monroe, J. M., Diener, M. J., Fowler, J. C., Sexton, J. E., & Hilsenroth, M. J. (2013). Criterion validity of the Rorschach Mutuality of Autonomy (MOA) scale: A meta-analytic review. *Psychoanalytic Psychology, 30*, 535-566. doi: 10.1037/a0033290

Newman, M. E. J (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics, 46*, 323-351. doi: 10.1080/00107510500052444

Nezworski, M. T., & Wood, J. M. (1995). Narcissism in the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice, 2*, 179-199. doi: 10.1111/j.1468-2850.1995.tb00038.x

Parisi, S., Pes, P., & Cicioni, R. (2005). *Tavole di localizzazione Rorschach, Volgari ed R+ statistiche [Rorschach location tables, Popular and R+ statistics]*. Disponibili Presso l'lstituto.

Parker, K. C., Hanson, R. K., & Hunsley, J. (1988). MMPI, Rorschach, and WAIS: A meta-analytic comparison of reliability, stability, and validity. *Psychological Bulletin, 103*, 367-373. doi: 10.1037/0033-2909.103.3.367

Pineda, J. A., Giromini, L., Porcelli, P., Parolin, L., & Viglione, D. J. (2011). Mu suppression and Human Movement responses to the Rorschach test. *NeuroReport, 22*, 223-226. doi:10.1097/WNR.0b013e328344f45c

Porcelli, P., & Kleiger, J. H. (2016). The "feeling of movement": Notes on the Rorschach Human Movement response. *Journal of Personality Assessment, 98*, 124-134. doi:10.1080/00223891.2015.1102146

Porcelli, P., Giromini, L., Parolin, L., Pineda, J. A., & Viglione, D. J. (2013). Mirroring activity in the brain and movement

determinant in the Rorschach test. *Journal of Personality Assessment, 95*, 444-456. doi:10.1080/00223891.2013.775136

Rorschach Archives. http://www.unibe.ch/university/services/-university_library/research/special_collections/rorschach_archives_and_collection/index_eng.html

Rosenthal, R., Hiller, J. B., Bornstein, R. F., Berry, D. T. R., & Brunell-Neuleib, S. (2001). Meta-analytic methods, the Rorschach, and the MMPI. *Psychological Assessment, 13*, 449-451. doi: 10.1037/1040-3590.13.4.449

Searls, D. (2017). *The inkblots: Hermann Rorschach, his iconic test, and the power of seeing*. New York: Crown/Random House.

Shaffer, T. W., Erdberg, P., & Haroian, J. (1999). Current nonpatient data for the Rorschach, WAIS—R, and MMPI-2. *Journal of Personality Assessment, 73*, 305-316. doi: 10.1207/S15327752JPA7302_8

Shaffer, T.W., Erdberg, P., & Meyer, G. J. (Eds.) (2007). International reference samples for the Rorschach Comprehensive System [Special issue]. *Journal of Personality Assessment, 89 (Suppl. 1)*.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112*, 140-154. doi:10.1037/0033-2909.112.1.140

Su, W.-S., Viglione, D. J., Green, E. E., Tam, W.-C. C., Su, J.-A., & Chang, Y.-T. (2015). Cultural and linguistic adaptability of the Rorschach Performance Assessment System as a measure of psychotic characteristics and severity of mental disturbance in Taiwan. *Psychological Assessment, 27*, 1273-1285. doi: 10.1037/pas0000144

Takahashi, M., Takahashi, Y., & Nishio, H. (2009). ロールシャッハ・テスト形態水準表 [*Rorschach Form Quality table*]. Tokyo: Kongo Shuppan.

Tuk, M. A., Zhang, K., & Sweldens, S. (2015). The propagation of self-control: Self-control in one domain simultaneously improves self-control in other domains. *Journal of Experimental Psychology, 144*, 639-654. doi: 10.1037/xge0000065

Viglione, D. J., & Giromini, L. (2016). The effects of using the International versus Comprehensive System norms for children, adolescents, and adults. *Journal of Personality Assessment, 98*, 391-397. doi: 10.1080/00223891.2015.1136313

Viglione, D. J., & Hilsenroth, M. J. (2001). The Rorschach: Facts, fiction, and future. *Psychological Assessment*, *13*, 452-471. doi: 10.1037/1040-3590.13.4.452

Viglione, D. J., & Rivera, B. (2013). Performance assessment of personality and psychopathology. In J. R. Graham, J. A. Naglieri, & I. B. Weiner (Eds.), *Handbook of psychology, Vol. 10. Assessment psychology (2nd ed.)* (pp. 600-621). Hoboken, NJ: John Wiley.

Villemor-Amaral, A. E., Yazigi, L., Nascimento, R. S. G. F., Primi, R., & Semer, N. L. (2007). *Localização, Qualidade Formal e Respostas Populares do Rorschach no SC em uma Amostra Brasileira [Location, Form Quality and Popular Responses in the Rorschach CS in a Brazilian sample*]. Em: III Congresso Brasileiro de Avaliação Psicológica. João Pessoa, PB, Brasil.

Wood, J. M., Garb, H. N., Nezworski, M. T., Lilienfeld, S. O., & Duke, M. C. (2015). A second look at the validity of widely used Rorschach indices: Comment on Mihura, Meyer, Dumitrascu, and Bombel (2013). *Psychological Bulletin, 141*, 236-249. doi: 10.1037/a0036005

Wood, J. M., Lilienfeld, S. O., Garb, H. N., & Nezworski, M. T. (2000). "The Rorschach test in clinical diagnosis": A critical review, with a backward look at Garfield (1947). *Journal of Clinical Psychology, 56*, 395-430. doi: 10.1002/(SICI)1097-4679(200003)56:3<395::AID-JCLP15>3.0.CO;2-O

Wood, J. M., Nezworski, M. T., & Garb, H. N. (2003). What's Right with the Rorschach? *The Scientific Review of Mental Health Practice, 2*, 142-146.

Wood, J. M., Nezworski, M. T., & Stejskal, W. J. (1996). The comprehensive system for the Rorschach: A critical examination. *Psychological Science, 7*, 3-10. doi: 10.1111/j.1467-9280.1996.tb00658.x

Wood, J. M., Nezworski, M. T., Garb, H. N., & Lilienfeld, S. O. (2001). The misperception of psychopathology: Problems with norms of the Comprehensive System for the Rorschach. *Clinical Psychology: Science and Practice, 8*, 350-373. doi: 10.1093/clipsy/8.3.350

Wood, J. M., Nezworski, M. T., Lilienfeld, S. O., & Garb, H. N. (2003). *What's wrong with the Rorschach?: Science confronts the controversial inkblot test*. San Francisco: Jossey-Bass.

Zipf. https://en.wikipedia.org/wiki/George_Kingsley_Zipf.

Received August 4, 2017
Accepted August 10, 2017