The Centrality of the Response Process in Construct Validity: An Illustration via the Rorschach Space Response

Joni L. Mihura, Nicolae Dumitrascu, Manali Roy & Gregory J. Meyer


To link to this article: http://dx.doi.org/10.1080/00223891.2017.1306781

Published online: 27 Apr 2017.
The Centrality of the Response Process in Construct Validity: An Illustration via the Rorschach Space Response

Joni L. Mihura,¹ Nicolae Dumitrascu,² Manali Roy,¹ and Gregory J. Meyer¹

¹Department of Psychology, University of Toledo; ²The Danielsen Institute, Boston University

ABSTRACT
Recently, psychologists have emphasized the response process—that is, the psychological operations and behaviors that lead to test scores—when designing psychological tests, interpreting their results, and refining their validity. To illustrate the centrality of the response process in construct validity and test interpretation, we provide a historical, conceptual, and empirical review of the main uses of the background white space of the Rorschach cards, called space reversal (SR) and space integration (SI) in the Rorschach Performance Assessment System. We show how SR and SI’s unique response processes result in different interpretations, and that reviewing their literatures with these distinct interpretations in mind produces the expected patterns of convergent and discriminant validity. That is, SR was uniquely related to measures of oppositionality; SI was uniquely related to measures of cognitive complexity; and both SR and SI were related to measures of creativity. Our review further suggests that the Comprehensive System use of a single space code for all uses of white space likely led to its lack of meta-analytic support as a measure of oppositionality (Mihura, Meyer, Dumitrascu, & Bombel, 2013). We close by discussing the use of the response process to improve test interpretation, develop better measures, and advance the design of research.

Historically, many professionals and laypersons alike have regarded the Rorschach inkblot task (Rorschach, 1921/1942) with skepticism (Crews, 2004; Goode, 2001; Jensen, 1965). Why? A simple explanation could be that how the Rorschach works is rarely discussed. The link between the psychological operations in the response process (Schachtel, 1966) and the interpretation of the related score is rarely explained in the contemporary psychology literature (Exner, 2003; Exner & Erdberg, 2005). Consequently, some criticisms in the popular press have referred to Rorschach interpretation as being like astrology, palm reading, and tea leaf reading (Wood, Nezworski, Lilienfeld, & Garb, 2003).

Although mystery about how the Rorschach works can be an intriguing feature of the test, and historically some clinicians have capitalized on this potential mystery by using what Weiner (2003) called the "Ouija board" approach to interpretation—the related opaqueness can lead to mistrust and doubt by professionals and the public. Instead, a framework for conceptualizing how the test works should be grounded in reasonable and documented mechanisms regarding how the mind works. A disjunction between a Rorschach response and the resulting interpretation of personality features makes the test difficult to teach and also challenging to discuss its results with clients in a meaningful and empowering way. Although it is conceivable that some psychologists might be motivated to keep the Rorschach shrouded in mystery, this opaqueness can also be propagated inadvertently—simply because the direct links between the Rorschach variable’s coding and its interpretation were never taught to the psychologist. This omission in training could occur for various reasons, perhaps because the link was lost over time or it was never fully explicated in the first place.

An illustration of how the link between a variable’s coding and its interpretation can be lost over time can be shown with the Rorschach Comprehensive System’s (CS; Exner, 2003) Pair score. A Pair is coded when the respondent reports seeing two objects on each side of the card. Pairs are included in the CS Egocentricity Index along with the Reflection score, which is coded when a person sees a reflection in the symmetrical “sides” of the inkblot (e.g., "a person looking at their reflection in the mirror"). The CS interprets Reflections as an indicator of narcissism, a term derived from Narcissus in Greek mythology who falls in love with his reflection. Psychoanalytic theory subsequently used the term narcissism to denote autoeroticism (Freud, 1914). But how do Pair responses fit this interpretation when Pairs are not a reflection, but two separate objects? For Exner, that interpretive link did initially exist. In a 1969 article, in which Exner explored the relationship between Rorschach responses and narcissism, he stated, "[Pair] responses were included in the analysis in that they represent a type of mirror image although clearly in a much more subtle form" (Exner, 1969a, p. 327). However, over time, Exner’s conceptual link between the Pair response and egocentricity was lost because it was not discussed in his CS test manuals.
In contrast, Urist’s (1977) Mutuality of Autonomy scale (MA)—an empirically supported psychoanalytic scale (Graceffo, Mihura, & Meyer, 2014) depicting levels of development ranging from engulfing and overwhelming (Level 7) to mutual and autonomous (Level 1)—makes this link more explicit. MA Level 4, representing the reflection-mirroring stage, is coded when “one figure is seen as the reflection, or imprint, of another” (Urist, 1977, p. 5). Holaday and Sparks (2001) further elaborated Urist’s coding of pairs in Level 4 by emphasizing that the pairs must be described as indistinguishable, by using descriptors like “identical,” “twins,” or “exactly the same,” which fits Kohut’s (1971) conceptualization of narcissism that includes mirroring and twinnship. Without this elaboration of sameness, pairs of objects do not fit the theoretical construct of narcissism. In fact, it appears that Exner (1969a) might have followed this more restrictive rule for coding Pair responses than in his final CS coding guidelines, because the Pair means he reports in that article are substantiated by using descriptors like “identical,” “twins,” or “exactly the same,” which fits Kohut’s (1971) conceptualization of narcissism that includes mirroring and twinnship. Without this elaboration of sameness, pairs of objects do not fit the theoretical construct of narcissism. In fact, it appears that Exner (1969a) might have followed this more restrictive rule for coding Pair responses than in his final CS coding guidelines, because the Pair means he reports in that article are substantially lower than his nonpatient CS means.¹

Another way that the link between a Rorschach variable’s coding and its interpretation is unclear is when the original link was “discovered” by chance or through a strict empirical method of scale construction. The Reflection score also serves as a good example of a CS variable for which a finding was initially discovered by chance. As Exner (1969a) stated, “Something by accident, it was noted during data analysis for another Rorschach study, that a relatively high frequency of “reflection” responses, occurred in Rorschach protocols of overt homosexuals.”² (p. 325). Exner (2003) used exploratory statistical methods to develop his six CS constellation indexes, using a method of test construction similar to that used to create the original Minnesota Multiphasic Personality Inventory (MMPI) scales (the “traditional scales”) that relies strictly on empirical methods to choose the items (Hathaway & McKinley, 1940). That is, scale items were not chosen based on their content or their conceptual link to a personality characteristic, but based solely on their statistical ability to differentiate target and control samples. The empirical method of item selection starts with a large number of items, which are administered to the two groups (e.g., depressed vs. nonpatients) and statistically winnowed down to the best discriminatory items.

The emergence of a new Rorschach system (Exner, 1974) that emphasized the empirical method of variable selection and deemphasized theory was similar to and synchronous with major transformations of the Diagnostic and Statistical Manual of Mental Disorders (DSM; American Psychiatric Association, 1968, 1980). The CS was introduced as a new empirically based atheoretical Rorschach system during a time when serious empirical challenges were made to the reliability and accuracy of the DSM–II (American Psychiatric Association, 1968). The first edition of the CS test manual was published in 1974 (Exner, 1974), the same year that Spitzer and Fleiss (1974) published their seminal paper challenging the diagnostic reliability of the DSM–II. Essentially, Exner was to the Rorschach what Spitzer was to the DSM, in that they each helped usher in a new atheoretical, empirically focused version of the Rorschach and the DSM, respectively: the Rorschach CS (Exner, 1974) and the DSM–III (American Psychiatric Association, 1980). By doing so, Exner helped bring respect for the Rorschach as a scientific instrument and dissociate the test from the disparaged psychoanalytic theory.

However, as many scholars have noted, strict reliance on empiricism in test development—referred to as “dustbowl empiricism” (Butcher, 2000)—can come at a cost. Statistically, a strict empirical method of item selection capitalizes on chance. Therefore, the individual items, not simply the test scales themselves, must be cross-validated. Failure to cross-validate item selection was a criticism of the strict empirical method of test construction used to develop the MMPI (Helmes & Reddon, 1993) and the Rorschach CS constellation indexes (Wood, Nezworski, & Stejskal, 1996). A strict empirical method of item selection can result in items with low face validity—that is, it might not be readily apparent how an item is assessing its scale’s construct (and, indeed, it might not be a valid measure). Using the MMPI–2 (Butcher, Graham, Tellegen, & Kaemmer, 1989) as an example, items with low face validity were referred to as “subtle items” and items with a clear link to pathology were called “obvious items” (Hollrah, Schottmann, Scott, & Brunetti, 1995). Due to the lack of cross-validation of MMPI items selected for the traditional scales, many subtle items might be on their scales simply by chance (e.g., Weed, Ben-Porath, & Butcher, 1990).³

In sharp contrast to dustbowl empiricism, a more contemporaneous view of test development and construct validity (Borsboom, Mellenbergh, & van Heerden, 2003, 2004) argues that theory and an understanding of the response process that leads to the score should be primary in developing scales. Borsboom et al. (2004) compellingly argued that:

> The concept of validity may never have been necessary because the instruments were generally set up on the basis of an idea of how they would work. In that case, the question of what it is, precisely, that is measured can simply be resolved by pointing to the processes that lead to the measurement outcomes. (p. 1067)

Regarding our definition of the response process, as described in Meyer, Viglione, Mihura, Erard, and Erdberg (2011):

> [The response process refers to] the factors that lead to or produce the task behaviors captured by a particular code. They are the psychological elements that are present in the process of generating a response with a particular set of coded attributes. The relevant factors or psychological elements are embedded in the respondent’s coded behavior and imagery, which include his or her abilities, organizational efforts, styles of processing, feelings, ideas, motives, and conflicts. These factors include longstanding personality

---

¹In Exner (1969a), the Pair means were 1.95 and 2.61 (SD = 1.32 and 1.56), which are notably lower than nonpatient means using the standard CS coding guidelines: Exner (1974) = 7.3 (no SD given), Exner (2003) = 8.52 (SD = 2.18), and Exner (2007) = 8.82 (SD = 3.08).

²At the time Exner wrote this article, homosexuality was theoretically due to auto-eroticism or narcissism (see Exner, 1969a).

³Joni L. Hollrah is the same person as the first author, Joni L. Mihura.

⁴On the other hand, subtle items exist for which the link to the scale’s construct becomes clear after reflection, and are therefore less likely to be error. For example, although it might not be obvious to a layperson how answering false to “I like to flirt” belongs on MMPI Scale 2 (Depression), with just a little reflection a psychologist could connect this experience with the loss of libido seen in depressed patients.
characteristics, current state-like circumstances, and reactions present in the context of being assessed by a particular examiner at a particular time. (p. 330)

Therefore, our use of the term response process specifically targets the link between the psychological operations and behaviors that occur while giving a response and the associated interpretation as applied to the case. We discuss the response process further in subsequent sections, using the experiential knowledge gained by understanding the role of the response process in the interpretation of how respondents might integrate the background white space of the Rorschach card into their responses.

The purpose of this article is to illustrate how validity can be improved by conceptually realigning constructs or interpretive inferences with the response processes that lead to measurement outcomes. To illustrate the importance of aligning the response process with a measure’s construct, we focus on different ways a Rorschach respondent can use the inkblot’s background white space to produce a response. First, we review the history of different codes that have been used to designate the use of white space to show how their unique response processes differentially align with distinct interpretations. Next, we review the empirical literature for these codes to evaluate the convergent and discriminant validity findings that would be expected based on these distinct response processes. We close with a general discussion of how to link response processes with test interpretation to improve interpretive accuracy, develop better measures, refine existing ones, and advance the design of psychological research.

How the lack of empirical support for the CS Space variable led to a sharper focus on its response process

The CS test manual (Exner, 2003) cites five studies that support the validity of the CS Space variable. In contrast, the construct validity meta-analyses by Mihura, Meyer, Dumitrascu, and Bombel (2013) reviewed a more comprehensive set of studies and did not find support for the CS Space variable as a measure of oppositionality, either the overt behavior or the associated emotion (anger). In fact, the effect size was essentially zero for both measures of oppositionality, either the overt behavior or the associated interpretation as applied to the case. We discuss the response process further in subsequent sections, using the experiential knowledge gained by understanding the role of the response process in the interpretation of how respondents might integrate the background white space of the Rorschach card into their responses.

The CS test manual (Exner, 2003) cites five studies that support the validity of the CS Space variable. In contrast, the construct validity meta-analyses by Mihura, Meyer, Dumitrascu, and Bombel (2013) reviewed a more comprehensive set of studies and did not find support for the CS Space variable as a measure of oppositionality, either the overt behavior or the associated emotion (anger). In fact, the effect size was essentially zero for both measures of oppositionality, either the overt behavior or the associated interpretation as applied to the case. We discuss the response process further in subsequent sections, using the experiential knowledge gained by understanding the role of the response process in the interpretation of how respondents might integrate the background white space of the Rorschach card into their responses.

The CS test manual (Exner, 2003) cites five studies that support the validity of the CS Space variable. In contrast, the construct validity meta-analyses by Mihura, Meyer, Dumitrascu, and Bombel (2013) reviewed a more comprehensive set of studies and did not find support for the CS Space variable as a measure of oppositionality, either the overt behavior or the associated emotion (anger). In fact, the effect size was essentially zero for both measures of oppositionality, either the overt behavior or the associated interpretation as applied to the case. We discuss the response process further in subsequent sections, using the experiential knowledge gained by understanding the role of the response process in the interpretation of how respondents might integrate the background white space of the Rorschach card into their responses.

The purpose of this article is to illustrate how validity can be improved by conceptually realigning constructs or interpretive inferences with the response processes that lead to measurement outcomes. To illustrate the importance of aligning the response process with a measure’s construct, we focus on different ways a Rorschach respondent can use the inkblot’s background white space to produce a response. First, we review the history of different codes that have been used to designate the use of white space to show how their unique response processes differentially align with distinct interpretations. Next, we review the empirical literature for these codes to evaluate the convergent and discriminant validity findings that would be expected based on these distinct response processes. We close with a general discussion of how to link response processes with test interpretation to improve interpretive accuracy, develop better measures, refine existing ones, and advance the design of psychological research.

How the lack of empirical support for the CS Space variable led to a sharper focus on its response process

The CS test manual (Exner, 2003) cites five studies that support the validity of the CS Space variable. In contrast, the construct validity meta-analyses by Mihura, Meyer, Dumitrascu, and Bombel (2013) reviewed a more comprehensive set of studies and did not find support for the CS Space variable as a measure of oppositionality, either the overt behavior or the associated emotion (anger). In fact, the effect size was essentially zero for both measures of oppositionality, either the overt behavior or the associated interpretation as applied to the case. We discuss the response process further in subsequent sections, using the experiential knowledge gained by understanding the role of the response process in the interpretation of how respondents might integrate the background white space of the Rorschach card into their responses.

The purpose of this article is to illustrate how validity can be improved by conceptually realigning constructs or interpretive inferences with the response processes that lead to measurement outcomes. To illustrate the importance of aligning the response process with a measure’s construct, we focus on different ways a Rorschach respondent can use the inkblot’s background white space to produce a response. First, we review the history of different codes that have been used to designate the use of white space to show how their unique response processes differentially align with distinct interpretations. Next, we review the empirical literature for these codes to evaluate the convergent and discriminant validity findings that would be expected based on these distinct response processes. We close with a general discussion of how to link response processes with test interpretation to improve interpretive accuracy, develop better measures, refine existing ones, and advance the design of psychological research.

History of white space coding and the differences across Rorschach systems

Since its inception in Rorschach’s (1921/1942) Psychodiagnostik, the white space variable (S) has gone through several iterations. In Psychodiagnostik, Rorschach defined his S response as “Intermediate forms (S) [Zwischenformantworten] … in which the white spaces are interpreted rather than the black or colored parts of the figure [inkblot] which surround them” (Rorschach, 1921/1942, p. 39, italics added). Stated differently, “S is the symbol for interpretations in which the black parts of the figures are neglected in favor of the white spaces they outline” (p. 186). In fact, Rorschach deliberately provided prominent white space areas that would be available for responses on three cards (Cards II, VII, and IX; p. 52) and indicated any parallel series of inkblots should be similarly designed (p. 53). Rorschach noted that Space responses were “most common in stubborn, eccentric normals and in negativistic, scattered schizophrenics;” he believed that “S answers always indicate some sort of tendency to opposition” (p. 39).

Subsequently, all five early U.S. Rorschach systems included a version of this variable (Exner, 1969b). However, no two systems coded it exactly the same. Some differentiated types of white space use and some did not. Hertz (1970), Klopfer and Kelley (1942), and Rapaport et al. (1946) differentiated between white space use (a) used as the location for the main part of a person’s response (e.g., on Card VII, “a white lamp” to the DS7 area), and (b) used to embellish the main response (e.g., on Card I, “a fox’s head” with the white spaces at DdS26 as “its eyes and teeth”). Figure 1 provides a non-Rorschach inkblot illustration. The bottom, central white space can be seen as the form of a teddy bear with the top two pinkish-red spots as the...
inside of its ears and the two pinkish-red spots below as its eyes. The figure–ground reversal that focuses on the white space to create the teddy bear would historically be coded “Primary” or “Main” Space. On this same inkblot, the lighter central pinkish-red area can be seen as a mask with the lower white spot as its mouth. The mouth on the response would be a “Secondary” or “Additional” Space response in the earlier Rorschach systems because the inkblot proper forms the mask and the white space, as the mouth, is an embellishment. As discussed later, the newer Rorschach Performance Assessment System (R–PAS; Meyer et al., 2011) calls these codes, respectively, Space Reversal and Space Integration, a convention we use in all subsequent sections.

Beck, Beck, Levitt, and Molish (1961) and Piotrowski (1957) did not differentiate between these two types of white space responses, although Piotrowski did say, “However, one might assign a value of a whole point to each of the rare pure S [i.e., Space Reversal], and a half point to each of the combined S [i.e., Space Integration]” (p. 95), but this was not his convention. To complicate things, Rorschach systems often used different codes for essentially the same variable. For example, when a small or rare white space area was used for the main response, Klopfert, Ainsworth, Klopfert, and Holt (1954) used the code “S,” Hertz (1946) used the code “s,” and Beck (1944) used the code “Dds.” Furthermore, over time, some Rorschach systematizers changed their coding for white space. For example, initially Klopfert’s system used the code s for responses provided to small or rare white space areas (Klopfert & Sender, 1936); but, subsequently, it coded all Space responses as S, although continuing to distinguish between Space Reversal and Space Integration responses in the interpretation (Klopfert et al., 1954; Klopfert & Davidson, 1962; Klopfert & Kelley, 1942). Fonda (1951) and Exner (1969b) provided more information about the differences in coding white space across the early popular U.S. Rorschach systems.

When Exner compiled the key aspects of existing Rorschach systems into the CS (Exner, 1974), he also included a white space variable. However, like Beck and Piotrowski, he did not differentiate between the different types. Instead, Exner (1974) decided that “The symbol S is included in the scoring for location whenever a white space area of the blot is included in the response” (p. 57). Instead of using formal codes to differentiate uses of white space, Exner (1974) recommended that the “type of S (common, uncommon, separate, or combined) can be noted … in the qualitative evaluation of the protocol” (p. 57). Yet subsequent CS manuals (e.g., Exner, 2003) omitted this recommendation. Therefore, the CS distinction between the Space responses formed by figure–ground reversals versus other uses of white space faded over time.

**Space interpretation based on response process**

The consequences of Exner’s (1974) decision to use just one Space code for all uses of white space becomes clear when one reviews the literature on previous Rorschach systems. Historically, many authors emphasized the link between the psychological maneuvers—or “response processes”—involved in giving a Space Reversal response and Rorschach’s claim that “S answers always indicate some sort of tendency to opposition” (Rorschach, 1921/1942, p. 39). The response process link of Space Reversal to oppositionality was highlighted by Rapaport et al. (1946) when they said, “But how is it that space responses seem to refer to an oppositional tendency? The first answer appears to lie in the common-sense consideration that a subject who interprets the white spaces is actually doing the reverse of what the instructions imply” (p. 177, italics added). Fonda (1977) similarly stated:

Although [Rorschach] made no reference to the concept of the reversal of figure and ground, it is clear that such was the perceptual process he had in mind as mainly responsible for these responses. … A person who produces S does indeed seem to be performing in a way contrary to instructions. Instead of interpreting the black or colored parts of the ink blot, as Rorschach quite naturally expected, the subject interprets a white space. (p. 114)

Fonda (1977) astutely noted that this Space Reversal response-process interpretation is consistent with the definition of “negativism”7 provided by Bleuler, with whom Rorschach trained and completed his dissertation. Bleuler (1911/1950) defined negativism and illustrated the construct of negativism with numerous examples, which are summarized here to provide a better understanding of the intended construct:

We subsume under the term, negativism, a number of symptoms which have the common characteristic that a reaction which would be expected in a positive sense occurs in the negative sense instead. The patients cannot or will not do what is expected of them (passive negativism); or they do the very opposite or, at least, something else than what is expected (active or contrary negativism).

When patients should be getting up, they want to stay in bed. When they are supposed to be in bed, they want to get up. … To “Good day,” they say, “good-bye.” They do their work all wrong; sew buttons on the wrong side of the clothes. They eat their soup with a fork and their desert with a soup spoon. They continually sit down in someone else’s place, enter every bed but their own. They call our children by their surnames (which they have picked up somehow) instead of by their nicknames. A hebephrenic is asked to saw some wood; he hails some small boards instead, then he has to bring small boards, which he puts on the wrong rack. He is supposed to go down the staircase but resists; then suddenly takes the whole flight in one great leap. (pp. 191–192)8

As Bleuler (1911/1950) explained regarding the etiology of negativism in schizophrenia:

In schizophrenia, the patient’s relationship to the outside world has been altered; it has on the whole become a hostile one. The patients live in their autistic worlds. Thus it can be shown in many cases that the patients consider all stimuli emanating from the outside world, which they cannot block off, as unpleasant disturbances. As a result of this, negativistic attitudes develop. (p. 442)

Fonda (1977) noted that Rorschach’s interpretation of the oppositional Space Reversal response process was “unequivocally pejorative.” However, Fonda described several alternative positive interpretations by considering the potential healthy functions of the figure–ground reversal process. These include Freud’s (1913/1959) concept of the “general instinct of

7 Schizophrenic negativism is not the same as negative symptoms in schizophrenia.

8 Bleuler’s (1911/1950) concept of negativism was not linked with the physical aggression that many Rorschach authors have associated with oppositionality, and which resulted in studies exploring the relationship between Space and aggression, violence, and psychopathy (see Mihura et al., 2013, p. 593).
mastery,” whose aim Hendrick (1943) described as “to control or alter a piece of the environment, an ego-alien situation, by the skillful use of perceptual, intellectual, and motor techniques” (p. 314). Fonda also invoked Angyal’s (1965) concept of autonomy in the white space figure–ground reversal and White’s (1959) role of competence as an aspect of autonomy.

The figure–ground reversal inherent in giving a Space Reversal response—or “doing the reverse of what the instructions imply”—is not part of the Space Integration response process, the main distinctiveness of which, by nature of its coding, is that the white space is integrated with the inkblot proper. Historically, the Space Integration variable received scant attention in the theoretical and empirical literature. The most relevant information appears to be Beck’s (1933) measure of cognitive organizational effort—the Z score—which includes instances in which the white space is integrated with the inkblot proper. Exner (1974, 2003) followed Beck’s convention of including the integrated Space response into his Z score. All of this suggests that producing a Space Integration response requires the ability to cognitively integrate and organize one’s perceptions.

**There are distinct response processes in each type of Space response**

Existing research is consistent with the view that the response processes involved in Space Reversal and Space Integration responses are unique and distinct. Bandura (1954b) and Nelson (1954) both found that the perceptual figure–ground reversal process—based on the rate of perceptual reversals of a Necker Cube and other reversible figures—was more characteristic of Space Reversal than Space Integration—that is, $r = .32$ versus $.19$ (Bandura, 1954b) and $.48$ versus $.06$ (Nelson, 1954). These findings support the idea that the figure–ground reversal process is more specific to the Space Reversal than the Space Integration variable.

Similarly, the very low associations between these two different types of Space variables suggest they are measuring unique constructs. Holding the number of Responses constant, Bandura (1954b) found that the partial correlation between Space Reversal and Space Integration was only $.08$. Using R–PAS coding guidelines for these variables (Meyer et al., 2011), and excluding responses in which both scores were present, Dumitrascu, Mihura, Meyer, and Onofrei (2011) found that the correlation between Space Reversal and Space Integration was only $.08$. One would not expect to see such low correlations if these two types of Space responses were targeting the same construct—especially because they are both derived from a performance-based monomethod measurement (Campbell & Fiske, 1959; Meyer, 1996, 1997).

**Previous validity reviews of the Rorschach Space variable**

We identified four Rorschach white space validity reviews—a chapter by Fonda (1977), an article by Frank (1993), the CS test manual (Exner, 2003), and Mihura et al.’s (2013) CS validity meta-analyses. None of these reviews differentiated between the different types of Space responses when formulating their conclusions. Fonda’s review initially emphasized the phenomenological difference between Space Reversal and Space Integration; however, his conclusions did not. He also included studies in which the type of Space variable was undefined. Frank’s literature review did not differentiate between Space Reversal and Space Integration, and he concluded that “there does not seem to be any consistency in the findings” (p. 1113). However, he suggested, it is “possible that S may have different meanings … whether space is used as a main response … or is an addendum to a main response” (p. 1113), for which he stated “a program of research is needed” (p. 1114).

Exner’s (2003) review of the validity literature supported the CS Space variable as a measure of oppositionality, which was in diametric contrast with Mihura et al.’s (2013) meta-analyses that found an association of essentially zero with relevant validity criteria ($r = .01$). However, as previously noted, Exner’s review did not include any studies that used the CS coding criteria for Space, which was required for Mihura et al.’s meta-analyses. This is crucial because $87.5\%$ of CS Space is due to Space Integration. Only one third of the CS Space responses consist of the figure–ground reversals with a response process that fits with an interpretation of oppositionality or negativism. Thus, it is understandable why Mihura et al.’s meta-analyses did not support CS Space as a measure of oppositionality.

**The R–PAS Space variables go back in time to capture the response process intended by Rorschach**

A revitalized understanding of the unique response processes for Space Reversal and Space Integration provides a new lens from which to view the validity literature. Applying a response process-focused approach, Dumitrascu, Mihura, and Meyer (2010) conducted a preliminary review of the early Space literature and found support for the oppositionality interpretation of the figure–ground reversal variable. This was intriguing because, from the beginning, Rorschach’s proposed interpretation of Space was based on this response process, which was lost when the two types of Space responses were combined.

Dumitrascu et al. (2011) investigated the convergent and discriminant validity of Space Integration and Space Reversal using level of education as a proxy for cognitive complexity in the Romanian CS normative sample CS (Dumitrascu, 2007), which provided a substantial range of education (i.e., 4–17 years). Dumitrascu et al. recoded protocols using a prepublication version of the R–PAS Space Reversal and Space Integration coding guidelines. The relationship with education was strong for Space Integration ($r = .48$) and moderate for Space Reversal ($r = .25$); however, when Space Reversal codes were limited to responses without Space Integration, its correlation with education dropped to .08. Therefore, the use of white space for integration or elaboration was related to education as a proxy for cognitive complexity, whereas the figure–ground reversal use of white space was not. Further, these two Space variables showed very small associations with each other (.08). Thus, these initial findings suggested that Space Reversal and Space Integration were assessing two different constructs.

---

9Space Integration responses outnumber Space Reversal responses almost 3:1 ($M = 2.8$ vs. 1.1; Meyer et al., 2011, p. 310).
The preliminary review and data in Dumitrascu et al. (2010; Dumitrascu et al., 2011), combined with the revitalized interpretation of Space responses via their conceptually based response processes, guided decisions in the most recent Rorschach system: R–PAS (Meyer et al., 2011). The construct validity meta-analyses by Mihura and colleagues (2013) played a central role in choosing which variables to include in R–PAS (Meyer, Viglione, & Mihura, 2017; Meyer et al., 2011). That review placed the CS Space variable in the “no validity support” category (see Mihura et al., 2013, Table 3, p. 570), so it was not included in R–PAS. Instead, R–PAS returned to the older literature distinguishing between the Space Reversal and Space Integration types of responses.

R–PAS named these scores Space Reversal (SR) and Space Integration (SI) to emphasize their inherent response processes (Meyer & Mihura, 2010; Meyer et al., 2011). SR was placed on Page 1 of the Summary Scores and Profiles results, which is for the variables with the best empirical support; SI was placed on Page 2 for more tentative interpretation because it had a clear response process interpretation but relatively little research. The R–PAS developers created guidelines so SR and SI could be reliably scored. The reliability reported for the R–PAS norms was conducted using an early version of the coding guidelines, and the disagreements that arose from that coding were used to revise the final coding guidelines (see Meyer et al., 2011, pp. 435–439). The reliability study relied on six blinded coders, consisting of two R–PAS developers—Gregory J. Meyer and Joni L. Mihura—and four of their graduate students: Nicolae Dumitrascu, Sandra Horn, Wei-Cheng Hsiao, and Joshua Eblin. A total of 60 nonpatient adult protocols were coded independently by five of the examiners, with the remaining examiner coding 50 of those records. The mean pairwise intraclass correlations (ICCs) across all six coders were .85 for SR and .88 for SI. For the two R–PAS developers (Mihura and Meyer), the ICCs were .89 for SR and .96 for SI.

Using a mixed nonpatient and patient sample of 50 child and adult records and draft versions of the R–PAS coding guidelines, two graduate students (Amy C. Blume-Marcovici and Heidi L. Miller) trained by another R–PAS developer—Donald J. Viglione—coded the records for all R–PAS variables (Viglione, Blume-Marcovici, Miller, Giromini, & Meyer, 2012). Their protocol-level ICCs for SR and SI were .91 and .86, respectively. Kivisalu, Lewey, Shaffer, and Canfield (2016) recently blinded coders an R–PAS interrater reliability study using student coders and a sample of 50 adult nonpatients. Response-level ICCs for two independent coders were .82 for SR and .85 for SI. Using Cicchetti’s (1994) interrater reliability guidelines—where poor agreement is < .40, fair agreement is .41 to .59, good agreement is .60 to .74, and excellent agreement is .75 to 1.00—these scoring reliability values for SR and SI are excellent.

**A conceptual review of the Rorschach Space response literature**

In subsequent sections, we review the research literature using methodology similar to that used in Dumitrascu, Mihura, and Meyer (2013) to provide a conceptual understanding of Space Reversal and Space Integration. We were interested in both convergent and discriminant validity (Campbell & Fiske, 1959).

**Method**

We reviewed the Space literature using a systematic methodology, which we believe is important when describing any body of literature, but particularly the Rorschach’s, which has come under especially harsh scrutiny over the past couple of decades (e.g., Wood et al., 2003). Therefore, we describe our review methodology in detail.

**Defining the Rorschach Space variables and their construct labels**

**Space variables**

We reviewed the empirical literature for studies that differentiated between white space coding based on figure–ground reversal and white space used as an elaboration or embellishment of a response to the inkblot proper. The variables that fit these categories were (a) Rorschach’s (1921/1942) Space, (b) Primary and Secondary Space (Hertz, 1970), (c) Main and Additional Space (Klopfner et al., 1954; Klopfner & Kelley, 1942), (d) SR and SI (R–PAS; Meyer et al., 2011), and (e) any study that modified its Space coding to fit either of these categories (typically based on Fonda’s (1951) article that distinguished between these two types of Space responses). Because the CS Space variable is mostly comprised of SI (87.5%; see Meyer et al., 2011, p. 310), we also were interested to see if it assessed cognitive complexity and so included it as a proxy for SI.

**Space Reversal and Space Integration constructs**

To choose relevant findings, we used the formal R–PAS SR and SI interpretations to target their core constructs (Meyer et al., 2011). R–PAS interprets SR within the Self and Other Representation domain, and states that it “may be prompted by creativity, individuality, oppositionality, or healthy self-assertive strivings” (Meyer et al., 2011, p. 362). We viewed the oppositionalism inherent in the SR response process as consonant with Bleuler’s (1911/1950) view of negativism, not as consonant with the overt aggression that characterizes oppositional defiant disorder in the *DSM–5* (American Psychiatric Association, 2013). Therefore, in contrast to Exner (2003), we concluded that physical aggression and psychopathy do not target the bullseye of SR’s core construct.

R–PAS locates SI within the Engagement and Cognitive Processing domain. Its interpretive postulates focus on the constructs of “integration” and “synthesis,” also stating that SI is “indicative of complex and flexible thinking” (Meyer et al., 2011, p. 367), to which the Response Process section adds “and possibly creativity” (p. 333). Therefore, for SI, we focused on findings with criterion variables that targeted these cognitive constructs.

We required the criterion variables for SR and SI to be behaviorally based. To define “behavior,” we used the definition provided by Baumeister, Vohs, and Funder (2007), which states:

A study qualified as having behavior if any element involved behavior—that is, if the study used any manipulation, any dependent measure, or even used behavior as the conduit for manipulating the independent variable (e.g., taking a test and getting feedback on it). … Self-reports of past behaviors or of hypothetical behaviors did not count. The use of archival behavioral data (e.g., crime statistics; donating blood) qualified. Reading about someone else’s behavior was not counted as behavior. (p. 399)
Literature search

We conducted PsycINFO and MEDLINE database searches using SR and SI construct keywords derived from the previously described R–PAS manual’s interpretations of these variables (Meyer et al., 2011, pp. 333, 362, 367). The final search to update our database was conducted on November 4, 2016. We initially used the search terms “Rorschach AND Space,” but many relevant articles did not use the term “Space” in their title or abstract. Therefore, we searched titles and abstracts of these databases using the search terms “Rorschach AND (Space OR opposition OR assertive OR individuality OR creative)” for SR and “Rorschach AND (cognitive OR complex OR creative OR integrative OR flexible OR synthetic)” for SI.

We limited the search to articles published in the English language with methodology classified as an empirical study. After deleting duplicates, there were 102 articles to review for SR and 356 articles to review for SI, for a total of 398 articles (an overlap of 60 articles). We also reviewed articles cited in the previously described Space reviews (Exner, 2003; Fonda, 1977; Frank, 1993) for potentially relevant findings, which added 12 articles to review for SR. For discriminant validity purposes, we included Mihura et al.’s (2013) meta-analytic findings for the CS Space variable in relationship to externally assessed criteria targeting oppositionality, which added nine articles with 13 findings.

We focused on the peer-reviewed journal literature with two exceptions. First, Charek, Meyer, and Mihura (2016) provided validity data for SI by documenting that it was under conditions of mental fatigue or ego depletion. In an as yet unpublished dissertation, Charek (2016) did not replicate these findings. Because two of the authors (Joni L. Mihura and Gregory J. Meyer) were on Charek’s dissertation committee and aware of these findings, we believed it was important to include this unsuccessful replication. Charek (2016) also reported the relationship between SR and SI with cognitive ability measures, so those findings are included as well. Second, we included the unpublished results from Dumitrescu et al. (2011) because that research was undertaken for this article. Therefore, in total, there were 412 studies to review for convergent validity findings—116 for SR and 356 for SI (with an overlap of 60 studies).

Selection procedure

The first author conducted the literature review, consulting with the second author for difficult decisions. We reviewed all relevant studies reported in three Space reviews (Exner, 2003; Fonda, 1977; Frank, 1993). For the PsycINFO and MEDLINE search, we did not obtain the full-text article if the title or abstract indicated that (a) the relevant Space variable was not being used (e.g., for SR, the CS Space response was being studied), or (b) the relevant constructs were not being assessed (e.g., for SI, the criterion variable only targeted oppositionality). The validity criterion had to clearly target the SR and SI characteristics described in the R–PAS manual and used in the search terms. If we could not reliably determine from the full-text source which type of Space variable was used, we excluded the study.

Study quality

To ensure better study quality, we excluded results that selectively reported significant findings or that contained errors for which (a) the correct value was not computable from the article, or (b) the study authors were not able to provide the correct data. Coding reliability was a study quality criterion. We did not expect Rorschach studies published before Weiner’s (1991) recommended minimum level (i.e., 80% agreement) to report coding reliability. Therefore, we only excluded findings that provided evidence of low reliability, not studies that did not report reliability. We used the same level of reliability for both Rorschach and criterion variables as used in Graceffo et al. (2014) of ICC < .60, Pearson r < .70, or percent agreement < 80% (Weiner, 1991).

Finally, we considered number of Responses (R) as a relevant study quality characteristic. The fact that people can provide a variable number of responses to the 10 inkbloks has been discussed as a potential problem for decades (Cronbach, 1949; Meyer, 1992). Although we did not generally require studies to control for R, the Beck method of Rorschach administration is an exception because it instructs respondents to “be sure to tell the examiner everything” (italics added) that you see on the card as you look at it” (Beck et al., 1961, p. 2). The Beck approach results in a large and variable number of Responses (e.g., see Katko, Meyer, Mihura, & Bombel, 2010, p. 595). Therefore, study findings were excluded if they used the Beck administration method and did not control for R.

Discriminant validity

Using the multitrait–multimethod approach to construct validity (Campbell & Fiske, 1959), we evaluated both convergent validity and discriminant validity. That is, if a study reported the association between one type of Space response and a measure assessing its targeted construct (e.g., SR and oppositionality), as well as the association between the other type of Space response and the same criterion measure (e.g., SI and oppositionality), we report each of these findings. By comparing only within-study results, we hold study design constant so that it is not a confounding variable.

Results

Description of sample

Our search strategy identified 23 independent studies with 25 findings that fit the convergent validity criteria and 17 findings that fit the discriminant validity criteria. Of these, 10 articles with 19 convergent validity criteria were identified in the PsycINFO and MEDLINE search strategy. Two additional articles were included from the published Space reviews (Exner, 2003; Fonda, 1977; Frank, 1993; Stein, 1973) and De Koninck and Crabbe-Declèves (1971) each contributed one convergent validity finding. Charek’s (2016) dissertation had three findings that fit the convergent validity criteria and three that fit the

---

10To be consistent with Exner’s (2003) interpretation of the Space response, Mihura et al.’s (2013) meta-analytic criteria for CS Space included conduct disorder, psychopathy, and physical aggression; however, as previously noted, this is not consistent with R–PAS’s view of SR’s response process interpretation.

11We included two articles (Bandura, 1954a; Stein, 1973) that used a Space variable that combined SR and SI responses because the methodology purposefully weighted the SR response heavily.
discriminant validity criteria. Dumitrascu et al. (2011) had one finding that fit the convergent validity criteria and one that fit the discriminant validity criteria. Nine studies with 13 discriminant validity findings were included from the Mihura et al. (2013) CS meta-analyses. In total, there were 14 SR studies with seven convergent validity findings and 13 discriminant validity findings. There were nine SI studies with 18 convergent validity findings and four discriminant validity findings. The reasons for study exclusions are described in subsequent sections.

Reasons for excluded studies

Space reversal
Ninety-nine studies were excluded from the SR PsycINFO and MEDLINE search because they were either (a) not investigating an SR variable \((n = 70)\); (b) using criterion variables that we judged as not targeting the bullseye of the SR construct, or were based on self-report or another inkblot variable \((n = 18)\); (c) not investigating Rorschach validity \((n = 8); \text{e.g.}, \text{reliability studies}); or (d) contained errors \((n = 1)\). Regarding the latter, Weltman and Wolfson (1964) was excluded because its summary table of information was incomplete and its mathematically nested cells could not be reconstructed from contradictory information contained in the text. All but one of the 18 studies whose criterion variables we judged as not targeting the bullseye were not designed to validate the Rorschach; instead, they used the Rorschach as a valid psychological test to understand how certain populations differed from control groups on various psychological characteristics (e.g., comparing controls to women who have had abortions, people who have chronic schizophrenia, or people who have nightmares). For the remaining study, we disagreed with the author about the suitability of the criterion as an indicator of oppositionality. Fonda (1951) considered it oppositional to answer “?” to a questionnaire with the response options of Yes, ?, and No. Although oppositionality could be one reason for responding ?, there are other possible reasons—such as pure uncertainty as to how to respond.

Regarding the 12 potential SR articles from the published Space validity reviews (Exner, 2003; Fonda, 1977; Frank, 1993) that were not identified in our PsycINFO and MEDLINE search,\(^{12}\) six were excluded because they did not use a figure–ground reversal coding for Space (Borgatta & Eschenbach, 1955; Counts & Mensh, 1950; Eschenbach & Borgatta, 1955; Lord, 1950; Murray, 1957; Ray, 1963). Four were excluded because we judged their criterion variables as not targeting the bullseye of SR’s construct. Rosen’s (1952) criterion was psychopathy and Schachtel’s (1951) was delinquency. As previously noted, we decided a priori that psychopathy and physical aggression do not target the bullseye of SR’s response process. Finn and Neuringer (1968) used left-handedness as a proxy for oppositionality, and there is no solid research to support this association. Fox and Blatt (1969) tested Rapaport, Gill, and Schafer’s (1945) hypothesis that scoring higher on Digits Backwards (DB) than Digits Forward (DF) might indicate negativity and a tendency to be oppositional. We had initially included their positive findings, although we did so cautiously due to the number of alternative reasons for why a person might score higher on DB than DF. The ability to mentally visualize digits in order and then reverse them might share psychological operations with reversing figure and ground in inkblot imagery. However, this is not encompassed by our target constructs for SR, and we eventually decided to exclude this finding when it was questioned by a reviewer. Finally, although we included Ingram (1954), we excluded several of his findings because they violated our interrater reliability criterion.

Space integration
Regarding the 349 excluded articles from the SI PsycINFO and MEDLINE search, based on the titles and abstracts, 67 were not investigating Rorschach scale validity (e.g., reliability, theory, Rorschach mentioned but not studied) and 91 did not include a relevant criterion variable—the latter included studies investigating sex differences, cultural differences, the relationship between the Rorschach and self-report measures, various medical conditions that are not expected to significantly affect cognitive abilities (e.g., hypospadias in childhood, myasthenia, Tourette’s, asthma), and other conditions not expected to differ on SI (e.g., dissociative identity disorder vs. borderline personality disorder, sex offenders vs. other offenders). The abstracts did not indicate that any of these studies were designed to study the Space variable. The remaining 191 articles were obtained for review, and 187 articles did not report SI results. Of the remaining four articles, two were not written in English. A third article did not hypothesize differences between groups on SI, but selectively reported results of several Rorschach variables that differentiated between their groups. The last article was excluded due to a data-reporting error that was unable to be corrected by the author (reporting an average proportion of 60% Space responses in a schizophrenia sample, which is implausibly high).

Orientation to the SR and SI construct validity results

The SR and SI convergent and discriminant validity findings are reported in Tables 1 and 2. Each table reports the study, sample size, study population(s), criterion variable, Rorschach system, coding reliability (if present), whether the analyses controlled for R, and columns for the convergent and discriminant validity effect sizes. As noted in the Method section, CS Space is reported as a proxy for SI.

Effect size calculation and reporting
All findings were converted to the same metric (Pearson’s \(r\)) for equivalent comparison across studies. When reporting SR and SI findings in Tables 1 and 2, all effects are reported as positive when in line with theoretical expectation and negative when in the opposite direction. On average, we expect cross-method associations \((r\) for Rorschach variables to be around .27 (see Mihura et al., 2013). Readers can use this benchmark when reviewing the results in Tables 1 and 2.

Main SR and SI construct validity results

Convergent validity results are in Table 1 for SR and Table 2 for SI. Discriminant validity results are reversed and in Table 2 for

\(^{12}\)We report these exclusions in detail because they have been included in previous Space reviews.
SR and Table 1 for SI. Overall, the SR and SI construct validity findings showed the expected convergent and discriminant validity effect sizes. For SR, six of the seven convergent validity findings are in the expected direction ($r = .14$–.74; Table 1). The one negative finding was from a study by Ingram (1954), who also observed one positive finding. College students with at least two SR responses, compared to those with no SR responses, expressed more hostility when their career potential was insulted during an interview; however, they expressed less hostility when completing a Tower of Hanoi task. Discriminant validity for SR, across four effect sizes, showed the expected lack of association with measures of cognitive complexity (Table 2). For SI (Table 2), all five convergent validity coefficients for the R–PAS SI variable are in the expected direction ($r = .14$–.48), and the 13 CS Space coefficients showed similar convergent validity ($r = .11$–.72). With respect to discriminant validity for SI, as would be expected, the CS Space variable showed no meaningful association with oppositionality criteria across 13 effect sizes (i.e., from Mihura et al.’s [2013] meta-analyses, shown in the first data row in Table 1).

**Table 1.** Space Reversal versus Space Integration relative to criteria of oppositionality, creativity, individuality, independence, or self-assertive strivings.

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Criterion variable</th>
<th>Rorschach system</th>
<th>Coding reliability</th>
<th>Control for R?</th>
<th>SR convergent validity ES</th>
<th>SI discriminant validity ES</th>
<th>CS S as SI proxy*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathological interpretation (oppositionality – the behavior or the affect)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mihura et al. (2013)</td>
<td>790</td>
<td>Results from CS meta-analysis</td>
<td>CS</td>
<td>Varies</td>
<td>Varies</td>
<td>.01 (k = 13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bandura (1954a)</td>
<td>59</td>
<td>Negativism, teacher ratings of high school students</td>
<td>Hertz, weighted</td>
<td>None</td>
<td>Yes</td>
<td>.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stein (1973)</td>
<td>40</td>
<td>Oppositionality, supervisor ratings of Naval enlisted men</td>
<td>Klopfer, weighted</td>
<td>None</td>
<td>Yes</td>
<td>.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingram (1954)</td>
<td>16</td>
<td>Hostility expressed: Insulting interview</td>
<td>Beck</td>
<td>None</td>
<td>No</td>
<td>.50*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Healthy interpretation (creativity, individuality, independence, or self-assertive strivings)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Koninck &amp; Crabbé-Declève (1971)</td>
<td>14</td>
<td>Field independence vs. dependence, rod-and-frame test (psychology undergraduates)</td>
<td>Klopfer</td>
<td>None</td>
<td>Yes</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bandura (1954a)</td>
<td>59</td>
<td>Assertiveness, teacher ratings of high school students</td>
<td>Hertz, weighted</td>
<td>None</td>
<td>Yes</td>
<td>.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ramachandra (1994)</td>
<td>60</td>
<td>Accomplished artists (creativity) vs. normal controls</td>
<td>Klopfer</td>
<td>None</td>
<td>Yes</td>
<td>.31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Column shown in italics represents discriminant validity. SR = Space Reversal; SI = Space Integration; ES = effect size, Pearson r; CS = Comprehensive System; k = number of effect sizes.

*Because the CS S score is largely composed of SI, it is used as its proxy. No studies investigated the actual SI score in relationship to oppositionality or the other criterion variables reported in this table.

**Discussion**

Rorschach scores that make use of the background white space were used to illustrate the role of the response process in construct validity. A historical and conceptual review illustrates how the coding of Space became disconnected from the original rationale for its interpretation. After realigning the interpretation of white space use with the response processes that led to its coding, a review of the research literature revealed the expected patterns of convergent and discriminant validity that support differential interpretations of SR and SI. That is, the SR response process aligns with the original rationale as a measure of oppositionality or negativism (Rorschach, 1921/1942) in that the examinee is “doing the reverse of what the instructions imply” (Rapaport et al., 1946, p. 177, italics added). In contrast, the SI response process does not align with the construct of oppositionality or negativity; instead, integrating the white background with the inkblot proper is consistent with complex, integrative, flexible, and possibly creative thinking.

**Systematic review of the Space Reversal and Space Integration literature**

Three previous narrative reviews of the Space literature exist (Exner, 2003; Fonda, 1977; Frank, 1993), in addition to the CS meta-analyses by Mihura et al. (2013). Ours is unique in both (a) differentiating between SR and SI types of white space coding, and (b) using a systematic review. The latter is important because 18 of the 20 articles mentioned in the previous narrative reviews were available for all three reviews (i.e., published in 1971 or before), yet there is minimal article overlap. No article is cited in all three reviews, and 12 of the 20 articles are cited in only one review. Although all studies in these three reviews were designed to target SR’s construct, almost half (9/20) did not use an SR type of criterion variable. Five of the 11 remaining studies, conducted half-century or more ago, used criterion variables that we judged as not targeting the bullseye of SR’s construct (psychopathy or delinquency) or as questionable measures of its construct (left-handedness, responding “?” vs. Yes or No) to a questionnaire, DB > DF. We excluded two other studies reported in these narrative Space reviews because one contained significant errors and the other used self-report criterion variables.

**Space Reversal**

Seven study findings fit our SR inclusion criteria. Two studies targeted the classic negativism interpretation of
oppositionality (i.e., doing the opposite as instructed, not the aggression inherent in oppositional defiant disorder). Bandura (1954a) found that high school students rated high on negativism by their teachers had significantly more SR responses ($r = .34$). Stein (1973) found that Navy enlisted men rated high on oppositionality by their immediate supervisors gave more SR responses ($r = .53$). Therefore, there are replicated findings for the relationship between SR and the oppositionality seen in negativism (Bleuler, 1911/1950). One study investigating the emotion related to oppositionality (hostility) found mixed results. Ingram (1954) found that SR was positively related to observer-rated hostility expressed during an insulting interview but negatively related (the opposite of the predicted direction) to observer-rated hostility while completing a challenging Tower of Hanoi type puzzle task. Ingram had defined these situations as interpersonally versus intellectually frustrating, respectively. Therefore, although this finding needs to be replicated (Ingram had a very small $n$ of 16), it is possible that the interpersonal and insulting dynamic of the interview pulled for a more negativistic response than the intellectually challenging puzzle.

Regarding the more uniformly healthy SR interpretations (creativity, individuality, independence, and self-assertive striving), there were three relevant findings, each targeting a different construct. De Koninck and Crabbé-Decleve (1971) found that Canadian college students who scored higher on a field independence measure (the rod-and-frame test) were more likely to produce SR responses ($r = .74$). However, this effect size is artificially inflated because the study used an extreme groups design with a very small sample size (i.e., out of 52

---

### Table 2. Space Integration versus Space Reversal relative to criteria of complex, flexible, synthetic, integrative, or creative thinking.

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Criterion variable</th>
<th>Rorschach system</th>
<th>Coding reliability</th>
<th>Control for R?</th>
<th>SI convergent validity $ES_r$</th>
<th>SR discriminant validity $ES_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>RI</td>
<td></td>
<td></td>
<td>SI proxy</td>
<td>SR w/o SI proxy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ferracuti, Cannoni, Burla, &amp; Lazzari (1999)</td>
<td>40</td>
<td>Creativity (Torrance)–Verbal</td>
<td>CS</td>
<td>None</td>
<td>No</td>
<td>.37</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fluidity</td>
<td></td>
<td></td>
<td></td>
<td>.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flexibility</td>
<td></td>
<td></td>
<td></td>
<td>.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Originality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Creativity (Torrance)–Figural</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fluidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flexibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Originality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elaboration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schwartz &amp; Canetti (2014)</td>
<td>40</td>
<td>Creativity (Remote Associations Test)</td>
<td>CS</td>
<td>ICC = .98</td>
<td>Yes</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(graduate and postgraduate students)</td>
<td>RI-PAS</td>
<td>ICC = .81</td>
<td>Yes</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td>Charek, Meyer, &amp; Mihura (2016)</td>
<td>97</td>
<td>Controls vs. ego depletion manipulation (undergraduates)</td>
<td>CS</td>
<td>ICC = .76</td>
<td>No</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>Charek (2016)</td>
<td>135</td>
<td>Controls vs. ego depletion manipulation</td>
<td>CS</td>
<td>ICC = .84</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Judgment of line orientation accuracy</td>
<td>R-PAS</td>
<td>ICC = .76</td>
<td>No</td>
<td>.41</td>
<td>-1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Delis–Kaplan Executive Function System (D–KEFS) Design Fluency total (undergraduates)</td>
<td>R-PAS</td>
<td>ICC = .76</td>
<td>No</td>
<td>.28</td>
<td>-.05</td>
</tr>
<tr>
<td>Lorio et al. (2010)</td>
<td>110</td>
<td>Healthy controls vs. progressive multiple sclerosis</td>
<td>CS</td>
<td>None</td>
<td>No</td>
<td>.32d</td>
<td></td>
</tr>
<tr>
<td>Tegtmeyer &amp; Gordon (1983)</td>
<td>38</td>
<td>Full-scale IQ (nonpatient children)</td>
<td>CS</td>
<td>None</td>
<td>Yes</td>
<td>.24</td>
<td></td>
</tr>
<tr>
<td>Di Nuovo, Buono, Colucci, &amp; Pellicciotta (2004)</td>
<td>46</td>
<td>Full-scale IQ (psychiatric &amp; rehabilitation inpatients)</td>
<td>CS</td>
<td>None</td>
<td>No</td>
<td>.35</td>
<td></td>
</tr>
<tr>
<td>Dumitrascu, Mihura, Meyer, &amp; Onofrei (2011)</td>
<td>111</td>
<td>Range of education (Grade 4–17, community volunteers)</td>
<td>R-PAS</td>
<td>ICC = .98</td>
<td>No</td>
<td>.48</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ICC = .94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rey–Osterrieth Complex Figure Delay</td>
<td>CS</td>
<td>ICC = .92</td>
<td>No</td>
<td>.31d</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Organization (child &amp; adolescent outpatients)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Column shown in italics represents discriminant validity. CS $S =$ Space coding; SI $=$ Space Integration; SR $=$ Space Reversal; $ES_r =$ effect size, Pearson $r$; CS = Comprehensive System; ICC = intraclass correlation; R-PAS = Rorschach Performance Assessment System.

*aBecause the CS $S$ score is largely composed of SI, it is used as its proxy.

*bLimited to SR responses on which SI was not also coded.

*cCohen’s $d$ for the CS space score reported in Lorio et al. (2010, Table 6) should be $-0.69$ not $-0.38$. We converted the correct value to $r$.

*dControlling for age and IQ.

---

13 As described in Bandura’s (1952) dissertation, in Bandura (1954a) “negativism” was rated on a 5-point scale ranging from “Stubborn and resistive, refuses to do things unless forced” at the high end to “Accepts discipline and orders without hesitation, docile and compliant” (p. 81) at the low end.

14 For Stein (1973), oppositionality was rated on three 7-point semantic differential scales with endpoints of cooperative–stubborn, agreeable–oppositional, and acquiescent–contrary.
tested, the seven highest and lowest scores on the rod-and-frame test), and should be replicated. Bandura (1954a) found a minimal relationship between SR and high school students rated as assertive by their teachers ($r = .14$). Finally, Rame-chandra (1994) found that accomplished Indian artists (those receiving prestigious awards) in music, literature, painting, sculpture, and dance produced more SR responses compared to normal controls.

**Space Integration**

In almost every instance, SI showed the expected relationships across a wide range of criterion variables that targeted complex, flexible, synthetic, integrative, or creative thinking, whereas SR was not related to these criteria. SI was positively related to different measures of creativity (Ferracuti, Cannoni, Burla, & Lazzari, 1999; Schwartz & Canetti, 2014), various neuropsychological measures (Charek, 2016; Smith, Bistis, Zahka, & Blais, 2007), IQ (Di Nuovo, Buono, Colucci, & Pellicciotta, 2004; Tegtmeier & Gordon, 1983), and education (Dumitrascu et al., 2011), and it was less common for persons with progressive multiple sclerosis compared to controls (Lorio et al., 2010). Charek et al. (2016) found that participants exposed to an experimental ego depletion manipulation had less complex protocols (e.g., fewer SIs) than controls, but the methodologically more complex follow-up study (Charek, 2016) did not. This negative finding in the SI review might be less a nonsupportive finding for SI and more the absence of an ephemeral ego depletion effect (e.g., Carter, Kohler, Forster, & McCullough, 2015), given that Charek (2016) documented clear validity for SI with cognitive performance criteria.

**Creativity and the Space response**

Interestingly, even though SR and SI were unrelated to each other statistically, they were each related to creativity. This might initially seem counterintuitive, but the SR and SI response processes likely capture unique aspects of creativity. For example, as shown in Table 1, Ferracuti et al. (1999) found that CS Space, which is largely comprised of SI, was related to all three types of verbal creativity (fluidity, flexibility, and originality), but to only two of the four types of figural creativity: figural flexibility and elaboration, not figural fluidity or originality. One would expect figural originality (i.e., the ability to create unusual, new, and different types of ideas with figures) to be more related to SR than to SI due to SR’s perceptual figure–ground reversal response process. Researchers should further investigate SR and SI’s unique associations with different types of creativity.

**Using the response process to improve the accuracy of test interpretation and empathic understanding of clients**

As previously noted, the R–PAS developers emphasized the response processes when deciding which Rorschach variables to include in the system (Meyer et al., 2011), and formal interpretation strongly emphasizes the response process. The R–PAS test manual provides two levels of interpretation: the first at the response level (the response process for individual coded variables, like Vista [V]), and the second at the protocol or summary level (integrating the response process of all coded variables in a score, like YTVC’).

In addition to focusing on the response process inherent in any particular test score, the interpreter must also consider the context, such as the inkblot, the instructions, the interpersonal situation, and the reason for the referral. For example, seeing numerous aggressive images on the Rorschach clearly indicates that aggressive images are on a person’s mind. However, it is fairly easy for people to notice and suppress these images (Benjestorf, Viglione, Lamb, & Giromini, 2013), which might occur when a person accused of a violent crime attempts to appear nonaggressive or when a person with obsessive–compulsive disorder has ego-dystonic intrusive violent images.

The interpreter should consider the interpersonal situation during the Rorschach administration and how similar processes might play out in analogous situations in everyday life. For example, the task of having to explain to the examiner why one perceives things the way one does on the Rorschach is like the everyday task of recognizing how one is feeling and reacting to experiences and being able to communicate that to another person. A respondent who has difficulty recognizing and articulating his or her experiences to the examiner in the testing situation is displaying an alexithymic process (an inability to describe the nuances and subtleties of one’s experiences or emotions). Accordingly, scores on Rorschach variables that reflect this process (e.g., FD, Blends, Form%) correspond to measures of alexithymia (Porcelli & Mihura, 2010).

When discussing a client’s Rorschach results with them, the clinician should stay as close as possible to the response process, and provide the client analogous examples as to how this process occurs in everyday life. Prioritizing the response process in an assessment enables a clinician to experientially empathize with clients through their test results compared to simply applying a score’s label. This approach to interpretation should resonate better with the client. Of course, the response process occurring during a person’s test performance is not usually entirely isomorphic to analogous behaviors in everyday life; therefore, the interpreter should be cognizant of the similarities and differences when applying test results to the client’s life. This caveat applies to all methods of assessment. As noted by Mihura and Meyer (in press), “On the WAIS–IV Block Design subtest, clinicians do not interpret the person’s ability to put together blocks in everyday life; they generalize the psychological operations that occur in the process of generating a response—e.g., visual analysis and synthesis in the case of Block Design—to similar operations in everyday life.”

To illustrate the differential interpretations of SR—including pathological versus nonpathological oppositionalism—we provide brief excerpts from three cases. We use a form of configurational (Bram & Peebles, 2014) and structural (Weiner, 2003) analysis, which considers the nature of the Space response and accompanying protocol features. We focus on the following major themes to help with the differential interpretations of SR responses. One key consideration with “doing the opposite of the instructions” in the SR figure–ground reversal is whether

---

15 As described in Bandura’s dissertation, in Bandura (1954a) “assertiveness” was rated on a 5-point scale ranging from “Insistent in making his opinions known, tries to do all of the talking, tends to force his views” to “Quiet and submissive, rarely or never asks questions or volunteers comments” (p. 81).
the intended target of the act is interpersonal. Using a response process perspective, interpersonal oppositionality can be ascertained from a Rorschach administration in two main ways: (a) when SR responses are accompanied with human content or human movement, and (b) the respondents’ interactions with the examiner are oppositional. *Pathological* oppositionality would be more likely if the responses contained a high degree of hostility and aggression and the respondent was hostile, aggressive, or uncooperative with the examiner. Pathological oppositionality would be less likely if the respondent was very cooperative, and expressed pleasure and interest in the Rorschach task. Finally, the degree to which movement scores that accompany the SR responses are passive or active could provide information as to how the oppositionality will be expressed. Subsequently, we provide three cases to illustrate how the response process varied in these key ways in a manner consistent with analogous situations observed in other parts of the assessment process or occurring outside the assessment.

**SR example Case #1: Max**

Kaakinen, Muzio, and Säävääli (in press) presented a violence risk assessment case conducted in Finland with Max, a 22-year-old visual communications student, due to vague threats he had made at school. Max has a history of two documented assaults (against a previous girlfriend who was breaking up with him, and a man who insulted him). School officials expressed fear of Max both as a potential aggressor toward others and himself. Throughout the assessment, Max was condescending and verbally aggressive. He easily lost his temper when asked questions that implied criticism (e.g., about his previous assaults). Max’s only assessment question was “I would really like to know why every time I try to express my artistic needs someone is calling the police!”

Max produced an extremely high number of SR responses (7, SS = 141) as well as an extremely high number of responses (R = 38, SS = 143). SR was still elevated after adjusting for protocol Complexity (SS = 141). Max’s behaviors during the Rorschach administration were hostile, oppositional, and uncooperative. For example, his high number of Rorschach responses can only be accomplished with R–PAS by “violating” the instructions to give “two … maybe three responses” per card and by not complying with the repeated reminders of the instructions. Max expressed irritation to the examiner each time she reminded him. Eventually she stopped giving reminders due to his irritation with her. Max’s SR responses were “a roguish face” (formed largely by only the white spaces on Card I), “a screaming face with a fearful, panic-like look,” “a torso of a body wearing a suit,” “a semi-surprised smiling face,” “an Aztec priest,” “a shock wave in an explosion,” and “a woman’s hipbone.” Hence, six of Max’s seven SR responses contain human contents, none of which are whole human representations, and five that were coded for Poor Human Representations. Max reported an extremely high number of aggressive images (AGC = 7, SS = 127). He had an equal number of passive and active movement scores with his SR responses. However, his two active movement responses were forceful and aggressive: the “screaming face” and the “explosion,” which Max elaborated as “a really powerful explosion that can be seen at a distance of hundreds of kilometers.” Therefore, the response processes accompanying Max’s protocol suggest an overt pathological oppositionality interpretation of his elevated SR, which is consistent with his behavior observed outside the Rorschach task.

**SR example Case #2: Tala**

Tala is a highly intelligent 24-year-old woman who presented for therapy with the first author after several failed therapies. Tala had been diagnosed with avoidant personality disorder, and she dropped out of a visual arts school due to depression. In the intake, Tala had several piercings—including a nose ring, lip ring, and numerous ear piercings—and her dress and attitudes were notably avant-garde. Tala had moved back in with her parents, rarely left her bedroom, creatively painted over her bedroom windows, and had not answered long-term friends’ phone calls or emails for over 2 years (about which she felt guilty). Tala summed up her motivation for therapy as “I want to have a life.”

Tala gave a high number of SR responses (6, SS = 141), as well as a very high number of Rorschach responses (R = 37, SS = 138). Like with Max, her SR was still elevated after adjusting for Complexity (SS = 141). However, unlike Max, Tala was cooperative during the Rorschach administration, after initially expressing concerns to her therapist about revealing her responses with someone she did not know. Also different than Max’s Rorschach administration, the examiner was using an early version of R–PAS R–Optimized administration, which did not consistently build in the reminder for “two … maybe three responses.” The examiner did, however, ask for the card back each time Tala gave four responses. Therefore, Tala might have begun to think that four responses per card was one of the rules.

Tala’s SR responses were “a wedding dress,” “a praying person wearing a robe,” “people with their backs together looking at papers,” “a cup in a museum from Ancient Greece,” “a tiny person on the horizon with snow on the ground,” and “the white face of a goat with a Fu Manchu mustache.” Hence, three of her six SR responses were whole humans, two of which had Good Human and one Poor Human Representation codes (the latter for low perceptual accuracy, FQ–). Tala had an average number of aggressive images (AGC = 3, SS = 100). Three of her SR responses were coded for movement and all were passive. Therefore, the response processes accompanying the SRs on Tala’s Rorschach protocol are notably different from Max’s. They do suggest that her oppositionality is interpersonally directed, but her protocol and her test behaviors are not hostile or aggressive, and she has evidence of healthy whole human representations.

These findings helped the therapist formulate and test the hypothesis that an underlying passive resistance might be driving Tala’s current state of living at home yet rarely coming out of her room as a compromise formation to individuate from her family. Her avant-garde presentation, attitudes, and as her therapist later learned, her artistic style, were likely also an expression of oppositionality. Although not as pathological as Max’s case, this oppositional style served as a compromise to simultaneously be different and get attention, yet keep most people somewhat distant.
SR example Case #3: Jill

Kamphuis, de Saeger, and Mihura (in press) presented a case conducted in the Netherlands of Jill, a 15-year-old female with posttraumatic stress due to chronic bullying at school and associated persistent depression. During the time she was being bullied, she was sexually assaulted by a same-age uncle. Jill had stopped attending school and was taking online classes. She was experiencing nightmares related to the bullying, engaging in self-harm, and, during one period of extreme stress, heard voices and believed she could communicate with plants. Jill’s parents described her as quiet and serious. She was not making progress in a psychotherapy using eye movement desensitization and reprocessing, and was referred for an assessment.

Jill produced three SR responses, which is elevated for her age (SS = 142). Jill’s SR responses were “a UFO in the night sky,” “the night sky … because the white spots look like stars,” and “a lamp with a shade.” Hence, none of her SR responses contained human representations or movement scores. She was very cooperative with the Rorschach administration, during which she stated that she liked to draw and sketch. She told the examiner that she enjoyed taking the Rorschach because it gave her the opportunity to “express [her] creativity.” Therefore, the expression of her figure-ground response process in everyday life might be less likely to be overtly interpersonal. Her SR responses, instead, hinted at a rather lonely and distant experience. Jill also produced a very elevated number of aggressive responses, instead, hinted at a rather lonely and distant experience. Her SR responses suggest that she is imaginatively seeking light in the darkness, or the middle white space on Card II (DS5) seen as a white spaceship flying through the night sky (D6).

Conclusion

Clinicians should closely attend to the response process with all assessment methods when making interpretations. For other Rorschach examples of using the response process in interpretation, see the cases discussed in Using the Rorschach Performance Assessment System (R–PAS; Mihura & Meyer, in press). For examples beyond the Rorschach (e.g., self-report instruments, clinical interviews, informant-ratings, narrative approaches like the Thematic Apperception Test, intelligence tests, implicit association tests), see Mihura and Graceffo (2014).

Align SI interpretation with its response process

Our main SI interpretive caveat is that the nature of the respondent’s SI responses fit the interpretation. Through a response process lens, the typical face response to Card I might not be the best SI example of “complex, flexible, synthetic, integrative, or creative thinking,” because the white spaces that comprise the facial features (DdS26) are critical bits that help form the percept rather than an elaboration to the main response, such as using the Card I white spaces as designs on a butterfly or moth. In contrast, less obvious uses of white space as facial features can indicate creative and flexible thinking, such as using DdS32 on Card VIII (inverted Whole) as the eyes on a mask. Other SIs that fit a more complex psychological operation occur when the inkblot proper is used to embellish or elaborate on a response in which SR constitutes the central object(s)—for example, the white spaces in Card I seen as “four ghosts dancing” and the inkblot incorporated as “… in the darkness,” or the middle white space on Card II (DS5) as a white spaceship flying through the night sky (D6).

Using the response process to develop better measures and refine existing ones

There are important ways that using the response process to conceptualize psychological test variables can result in better test validity research designs. As a general logical outcome of our major thesis—when choosing validity criteria for test validity studies, researchers should map the response process of the test variable onto the criterion variable, considering the ways in which they are similar and different. Returning to the Block Design response process example, note how its response process components are similar and dissimilar to the R–PAS Synthesis response. For example, their response process similarities include the ability to analyze and synthesize abstract visual information. Yet, Block Design is dissimilar from Synthesis in that it requires physical manipulation of the stimuli. The Synthesis response also seems to be the result of language-mediated, top-down mental processes, compared to the more bottom-up, visual-spatial mental processes activated in Block Design.

As another dissimilarity between Synthesis and Block Design, the Rorschach is a “typical performance test” (what a person will typically do when left to their own predilections) compared to IQ tests, which are “maximal performance tests” (what a person can do when asked to perform optimally to clear guidelines; Cronbach, 1960). R–PAS instructions do not ask the respondent to give responses that draw relationships
between the blot components (i.e., Synthesis). Therefore, a research design could use an experimental manipulation to bring the demand situations of the two tests closer together to test the components of their respective response processes. For example, if respondents were provided maximal performance instructions on the Rorschach specific to the Synthesis response process—for example, “Rather than reporting just one thing, your goal is to report two or more things in some relationship to each other”—respondents’ Synthesis scores might be more strongly related to their Block Design scores than with the traditional Rorschach instruction of “What might this be?”

Bornstein has written several articles and conducted experiments that capitalize on the response process by using experimental manipulations with a framework he refers to as the “process dissociation approach” (e.g., Bornstein, 2002, 2011, 2012). The experimental manipulation in the previous example with WAIS Block Design and the R–PAS Synthesis score uses a process somewhat like his process dissociation approach. Bornstein described a three-step process to conducting process dissociation research: (a) determine the score’s underlying processes and the context variables that alter these processes, (b) turn these process-altering variables into experimental manipulations, and (c) interpret the outcome in the context of these processes as well as any limiting conditions.

As described in our introduction, Borsboom and colleagues (2003, 2004) espoused a view of construct validity similar to Bornstein’s (2002, 2011), although they did not articulate the validation steps as clearly. Conceptually, Borsboom et al. argued that theory and an understanding of the response process that leads to the score should be primary in developing test scales, and that this requires experimental manipulation to determine cause and effect. An illustration of applying the approach Borsboom espoused can be seen in the studies by Charek (Charek, 2016; Charek et al., 2016)—cited in Table 1 (this article)—that used experimental manipulations of ego depletion prior to administering the Rorschach and cognitive tests compared to the results of a control group. Test variables assessing cognitive complexity, such as SI, were hypothesized to be less common in the experimental group (in which induced depletion should reduce cognitive complexity) compared to controls.

In the past decade, there has been a rise in studies targeting the neurological processes underlying Rorschach scores. For example, findings from an eye-tracking study by Dauphin and Greene (2012) are particularly relevant to the Space response. That is, although Rorschach designed Card II with a “prominent intermediate figure” (DS5) and Card IX with a “definite intermediate figure” (DS8), he intentionally designed Card VII to primarily target the figure–ground reversal process. About Card VII, Rorschach (1921/1942) stated, “The essential part is the white intermediate figure, a rather obvious oil lamp, rather than the black figures. This Plate presents the converse of Plate V, in that normals rarely see the lamp while schizophrenics frequently do” (p. 52, italics added). Rorschach’s expectation that schizophrenics are more likely to report SRs to Card VII is consistent with, and possibly driven, by Blueler’s (1911/1950) view of schizophrenic negativism. Using a nonpatient sample, Dauphin and Greene (2012) found that Card VII was the card with the shortest initial saccade latency. Respondents’ initial Card VII focus was on the middle white space; yet, their focus very quickly shifted to the inkblot proper, suggesting that the white space is not perceived as the goal of the “What might this be?” instruction even though it is the most prominent component.

In the past few years, several studies have shown an association between the Human Movement response (M) and mirror neuron activity in the brain (Ando et al., 2015; Giromini, Porcelli, Viglione, Parolin, & Pineda, 2010). Mirror neurons fire when primes act, and when they observe the same action by others. The functions of the mirror neuron system have been implicated in the ability to mentalize and empathize, consistent with M’s interpretation. Several other recent brain-focused studies have supported Rorschach variables’ interpretations and are consistent with the presumed underlying response processes. For example, using fMRI, Jimura, Konishi, Asari, and Miyashita (2009) found an association between Achromatic Color (C) and medial frontal lobe activity when receiving negative feedback, consistent with the interpretation of C as internally focused negative emotion. Asari and colleagues found evidence that unusual perceptions (as in FQu) are associated with brain activity that suggests personal emotional processing is involved—that is, with the amygdala modulating the connections between the temporopolar regions and the anterior prefrontal region and the bilateral occipitotemporal regions (Asari et al., 2008, 2010a, 2010b). This is only a sampling of the existing Rorschach-focused brain-based research, which can and should be incorporated into our understanding of the response processes that underlie Rorschach variables.

We strongly encourage psychological test researchers to familiarize themselves with test validation frameworks that target the response process. A framework from which to start is provided by the expert consensus classification process used in the intelligence and achievement test literatures (Flanagan, Ortiz, Alfonso, & Mascolo, 2002, 2006; McGrew & Flanagan, 1998). Applying this model to the response process components of other psychological tests would require assessment experts with the relevant specialties (e.g., psychosis, stress reactions, personality disorders) and assessment methods (e.g., self-report, inkblot, storytelling tests) to provide judgments as to what response process components lead to the test score, how these components map onto the targeted psychological construct, and identify the limiting conditions.

Conclusions
Historically, the psychological test validity literature has almost exclusively focused on the associations between different test variables rather than an understanding of the psychological operations inherent in the respondent’s response process. However, in the past decade or so, there has been an emerging interest in creating, validating, and interpreting psychological tests through an understanding of their response processes (Bornstein, 2002, 2011; Borsboom et al., 2003, 2004; Meyer et al., 2016; Meyer et al., 2011). To illustrate the importance of the response process in test interpretation, we provided a historical, conceptual, and empirical review of the different types of Space responses to show how focusing on their unique response processes leads to a clearer and more accurate understanding of their psychological constructs. We encourage clinicians and
researchers to apply the response process conceptual framework to interpreting psychological test results, designing new measures, and refining existing ones. It is vital that psychologists more clearly and systematically understand the response processes inherent in different assessment methods.

**Disclosure**


**References**

References marked with an asterisk indicate studies included in the systematic review results.


Hertz, M. R. (1946). Frequency tables to be used in scoring the Rorschach inkblot test (3rd ed.). Cleveland, OH: Department of Psychology, Case Western Reserve University.


