JOURNAL OF PERSONALITY ASSESSMENT, 72(1), 1–35

ARTICLES

The Convergent Validity of MMPI and Rorschach Scales: An Extension Using Profile Scores to Define Response and Character Styles on Both Methods and a Reexamination of Simple Rorschach Response Frequency

Gregory J. Meyer
Department of Psychology
University of Alaska Anchorage

Past research indicated the convergence of Rorschach and MMPI scales may be a function of (a) simple Rorschach response frequency (*R*) or (b) complex response–character styles on both methods. In this study, new criteria were developed for defining the second assumption using *F* and *K* from the MMPI and *R* and *Lambda* from the Rorschach. Although substantially different from the factor criteria used previously (κs = .45 and .30), the new criteria still produced the expected pattern of correlations among MMPI and Rorschach scales. Averaged across 17 constructs, the new criteria produced strong validity coefficients for patients with similar styles (*M* composite $r = .50$), though they were less effective for patients with discordant styles (*M* composite $r = -.27$). It was also demonstrated that *R* by itself does not moderate convergent validity. Rather, statistical modeling with two sets of 300 random samples (a) demonstrated the prior findings related to *R* were the result of sampling error and (b) supported the general hypothesis that Rorschach and MMPI scales correlate to the extent response–character styles correlate. Implications are considered.

The Rorschach and Minnesota Multiphasic Personality Inventory (MMPI/MMPI–2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) are the most frequently used instruments for the clinical assessment of personality. However, despite similar names, scales on these two tests generally are not correlated (Archer & Krishnamurthy, 1993a, 1993b; Meyer, 1997b). While this lack of association has prompted some to articulate ways in that the MMPI and Rorschach can compliment each other and profitably be used together in clinical evaluations (Finn, 1996; Ganellen, 1994, 1996; Meyer, 1996, 1997b), it has also raised questions about test validity (e.g., Archer & Krishnamurthy, 1997). In a thoughtful article aimed at furthering the scientific foundation for personality assessment, Archer (1996) recommended that any model designed to account for this lack of association should be both parsimonious and empirically testable.

Keeping in mind parsimony is relative to the complexity of the process under study, Meyer (1996, 1997b) proposed that Rorschach–MMPI independence is a function of at least three processes: (a) the inherent complexity of personality, (b) the ability of the MMPI and Rorschach to quantify distinct realms of personality, and (c) the unique sources of methodological bias that interfere with accurate measurement using either method. With respect to the last point, the literature has proposed two methodological factors that may contribute to Rorschach and MMPI convergent validity. One is *R*, the number of responses in a Rorschach protocol. The second consists of broad stylistic qualities that affect scores on both methods.

The basic findings regarding *R* were serendipitous (Meyer, 1993). Because patients can give a fluctuating number of responses to the Rorschach, it was thought that *R* may confound efforts to demonstrate cross-method validity. Consequently, it was hypothesized that similarly named MMPI–2 and Rorschach scales would be most strongly correlated when Rorschach protocols were of an average length. When scores were obtained from brief or long protocols, the scales were expected to be less related. However, the data did not fit this pattern. Scales were not correlated when Rorschach protocols were of an average length but they were positively correlated when the Rorschachs were lengthy and negatively correlated when they were short. *R* then seemed to reflect a willingness to acknowledge difficulties on a self-report instrument: When *R* was high, Rorschach and MMPI scales tended to agree; when *R* was low, the methods tended to disagree.

After considering these unexpected findings, it was subsequently proposed that *R* alone was not a prime factor for determining cross-method convergent validity (Meyer, 1994, 1997b). Rather, convergence was hypothesized to be a function of more general stylistic qualities that reflect how patients tend to approach or interact with each of the assessment methods. Psychometrically, these stylistic qualities are seen as sources of "systematic error"[1] in test data. When present, they

[1]Method specific variance is termed *systematic error* (as opposed to the random error quantified by reliability coefficients) because it limits convergent validity in nomothetic research. However, this "error" should not be equated with "meaningless" information. A clinician working idiographically with a patient can learn very useful information from test data indicating a patient is overly defended, magnifying problems, or unable to distance herself from the inkblot task.

cause observed scale scores to be higher (or lower) than what is genuinely accurate. In the literature such stylistic qualities often have been termed *response styles.* However, this term is frequently taken to mean the deliberate manipulation of test data (e.g., "fake bad" or "fake good"). Although deliberate distortions are important, the processes considered here also emerge from genuine elements of character structure (see Meyer, 1997b; Nichols & Greene, 1997). For instance, some patients over-endorse pathology because they genuinely see themselves as compromised or deficient across a host of psychological traits. Other patients under-endorse pathology because their defensive structures demand they perceive themselves from an overly idealized vantage point (e.g., Colvin, Block, & Funder, 1995; Robins & John, 1997; Shedler, Mayman, & Manis, 1993). Although somewhat cumbersome, to emphasize these stylistic patterns can be deliberate manipulations as well as manifestations of intrinsic character structures, they will be referred to as *response–character styles* (although *test interaction styles* also would be apt).

Response–character styles can emerge from qualities related to honesty, motivated goals, spontaneity, articulateness, dynamic defensive operations, insight, degree of self-concept differentiation, sensitivity to distress, and psychic boundaries, among other things. What unifies these diverse qualities is that they can all confound the accurate measurement of other Rorschach or MMPI constructs. Consider spontaneity. At times one may wish to measure this trait directly. If so, then a patient's degree of behavioral spontaneity when completing the MMPI or Rorschach may actually assist accurate measurement of the construct. However, almost every MMPI or Rorschach scale reflects a construct thought to be measured regardless of spontaneity. For instance, it is expected that scores on Scale 2 from the MMPI–2 or Rorschach Depression Index (*DEPI*) will reveal something about the patient's degree of depression regardless of how spontaneous the patient was when completing the testing task.[2] Unfortunately, this is not so.

The extent to which a patient is spontaneous or inhibited when articulating Rorschach responses helps determine the overall complexity or "richness" of the pro-

[2]In practice, sophisticated clinicians never have such global expectations. Thus, even though interpretive textbooks (e.g., Archer, 1992; Butcher & Williams, 1992; Exner, 1993; Graham, 1993; Greene, 1991) provide generic interpretations for specific scale elevations, ultimately every test score is interpreted in light of other test scores, history information, and observed behavior. Researchers also do not really believe that a given test scale measures its intended construct with unerring accuracy. However, almost without exception, assessment research is conducted as if test scales accurately measure their intended constructs—regardless of the bias and error that may be influencing each person's score. In this respect, it is almost universally "expected" that test scales provide relevant information regardless of response–character styles. However, it is also in this respect that our right hand seems not to know what our left hand is up to. Even though we truly know better, for some reason we treat nomothetic testing data as if they were impervious to the substantial influence of idiographic contextual factors.

tocol. In turn, this affects the absolute magnitude of most other scores, including those on the *DEPI* (see Meyer, 1992b, 1993). Similarly, the extent to which a patient is spontaneous or inhibited when answering MMPI items helps determine the suppression or elevation of all scales on an MMPI profile, including Scale *2* (see Edwards, 1957, 1964; Edwards & Edwards, 1991; Jackson, Fraboni, & Helmes, 1997). Thus, a patient's degree of spontaneity contributes to score elevations on a host of scales that were never designed to measure spontaneity. In turn, this impairs the ability of these scales to measure the constructs they were designed to quantify.

Table 1 provides a schematic review of some ways that response–character styles may be manifest on the MMPI and Rorschach. Five prototypical modes of interaction are presented for each of the testing tasks. Two relate to defensive operations and are hypothesized to produce a constricted yield of test information. Two others relate to exaggeration or excessive endorsement and engagement. They are hypothesized to produce a dilated yield of information. For each pattern, the table also presents scores that should be sensitive to these testing behaviors. Expectations are provided for the first unrotated principal component from each test, as well as for the *F* and *K* scales from the MMPI and *R* and *Lambda* from the Rorschach.

Several points can be noted about these hypothesized styles. First, from a measurement perspective, they can be considered sources of method variance (see Campbell & Fiske, 1959; Cole, Howard, & Maxwell, 1981; Cole, Truglio, & Peeke, 1997) because they are intrinsically tied to a particular method of assessment. Thus, even though the styles are arrayed in parallel columns and share some formal qualities, the actual personality characteristics that are associated with each style vary from method to method.[3]

Second, the styles are believed to have "fuzzy boundaries" because there is no clear point that separates an optimal style (Style 3) from one that will generate systematic error in other test scores (Styles 1, 2, 4, or 5). Third, there is nothing in the test data alone that can differentiate deliberate efforts to bias the test data (Styles 1 and 4) from characterological qualities that cause bias (Styles 2 and 5), because both can lead to an identical pattern of observed test scores. Thus, it is anticipated that clinicians and researchers will only be able to distin-

[3]To the extent that specific tests share common tasks or procedures for gathering information, stylistic method variance should be correlated across tests. Thus, stylistic method variance observed with a self-report inventory should correlate with stylistic variance observed during a structured interview, because both tasks depend on the deliberate and honest reporting of characteristics represented in the patient's conscious awareness (Alterman et al., 1996; Pogge, Stokes, Frank, Wong, & Harvey, 1997). Similarly, stylistic method variance observed on the Rorschach should correlate with stylistic variance observed on a cognitive performance task like the Wechsler scales (e.g., Glutting, Oakland, & Konold, 1994; Oakland & Glutting, 1990), because both tasks depend on the patient's behavioral investment with the task and active engagement with an examiner.

TABLE 1
Hypothesized Sources of Method Variance From Response–Character Styles on the MMPI–2 and the Rorschach

Style	*MMPI–2*	*Rorschach*
Deliberate defensiveness–constriction (Style 1)	Patient is situationally guarded, nondisclosing, and consciously responding to items in a socially desirable fashion (Style 1–M; low *FF;* higher *K*, lower *F*).	Patient is situationally constricted and leery about what information will be obtained from him or her; offers few responses, minimal task investment, resistive inquiry, and minimal elaboration (Style 1–R; low *FF;* lower R, higher *Lambda*).
Characterological defensiveness–constriction (Style 2)	Patient is characterologically defended, lacking insight, lacking self-awareness, and unable to accurately describe himself or herself. Portrays self as overly virtuous or ideal (Style 2–M; low *FF;* higher *K*, lower *F*).	Patient has genuine cognitive limitations, impoverished internal resources, or inhibitory defenses; patient is bewildered by the abstractness or complexity of the task (Style 2–R; low *FF;* lower *R*, higher *Lambda*).
Optimal range (Style 3)	Patient is open, insightful, and as forthcoming as possible about his or her problems or lack thereof (Style 3–M; moderately low to moderately high *FF*, *K*, and *F;* all correspond to extent of genuinely recognized pathology).	Patient is articulate, cooperative, spontaneous, and actively engaged; responses, determinants, and degree of synthesis are given in accordance with internal predilections (Style 3–R; moderately low to moderately high *FF*, *R*, and *Lambda*).
Deliberate exaggeration–dilation (Style 4)	Situational circumstances cause or motivate patient to be overly dramatic in his or her presentation and to portray self in an unrealistically pathological fashion (Style 4–M; high *FF;* higher *F*, lower *K*).	Situational circumstances cause or motivate patient to be overly dramatic, exhaustive, creative, complex, or shocking in his or her responses (Style 4–R; high *FF;* higher R, lower *Lambda*).
Characterological exaggeration–dilation (Style 5)	Patient is intrinsically fragile, self-critical, or hypersensitive to distress; honestly sees self as deficient and experiences self in an overly pathological and symptomatic light (Style 5–M; high *FF;* higher *F*, lower *K*).	Poor boundaries or pressing dynamics generate intense absorption with the task and create problems disengaging from the stimuli (Style 5–R; high *FF;* higher R, lower *Lambda*).

Note. Style X–M = MMPI style; Style X–R = Rorschach Style; *FF* = first factor (i.e., first unrotated principal component).

guish these styles after considering test data in light of the patient's history, observed behavior, and circumstances surrounding the evaluation. Fourth, these interaction styles are not mutually exclusive. Patients may have a genuine propensity to see themselves as fragile and deficient (Style 5–M) at the same time that they have some motivation to deliberately malinger disturbance (Style 4–M) or feign health (Style 1–M).

Finally, in general, observed scores on scales of symptomatology will be artificially inflated with Styles 4 and 5 and artificially deflated with Styles 1 and 2. However, not every test score will be affected to the same degree. Consider Style 2. On the MMPI, patients who are defended for characterological reasons will often elevate on Scales *K*, *L*, or *True Response Inconsistency* (with a "nay saying" bias) and suppress on Scale *F*. Elevations and suppressions on these validity scales will accurately reflect the patient's defensive character structure and they can be interpreted directly because they were designed to quantify such a style of interacting with the test. Some additional scales may remain impervious to this defensiveness. For instance, a patient's scores on Scale *5*, phobia-like Fears (*FRS*), or Bizarre Mentation (*BIZ*) may remain unaffected by the style because the patient's defensive structure may not require minimization of these issues. At the same time, defensiveness is likely to suppress scores on distress scales like Scales *2*, *4*, *7*, and so on, while elevating scores on the *Es* or *Re* scales. Observed scores on the latter cannot be trusted as genuine indications of the characteristic the scales were designed to measure because the patient's test taking approach artificially influenced the score (i.e., the *Es* scale becomes elevated because of defensiveness, not because the patient has genuinely enhanced ego resources).

The same considerations apply to Rorschach data. In the face of Style 2, observed scores on the Coping Deficit Index (CDI) could be interpreted with a fair degree of confidence because the CDI is intended to quantify such characterological limitations and it draws on specific scores sensitive to these qualities. However, in the presence of Style 2, one should be cautious about interpreting observed scores on many other Rorschach scales. For instance, the *WSum6* is likely to be artificially deflated secondary to the patient's limited engagement with the task.

Recent research has supported several points about these styles of interacting with the MMPI–2 and Rorschach. First, although these patterns are highly correlated within a method family (e.g., within the self-report method, $r = .85$), MMPI response–character styles are independent of Rorschach response–character styles, having an average correlation of .01 (Meyer, 1997b). Thus, the characteristics that make patients openly engaged or defensively limited when interacting with the MMPI are distinct from the characteristics that make patients openly engaged or defensively limited when interacting with the Rorschach. This also demonstrates that the behaviors listed in the Rorschach and MMPI–2 columns of Table 1 are empirically unrelated.

Second, Meyer (1997b) reported that carefully selected Rorschach and MMPI–2 scales of psychosis, affective distress, and interpersonal wariness were uncorrelated when response–character styles were ignored, having an average cross-method convergent validity coefficient of $r = .03$. At the same time, this study revealed strong positive validity coefficients when analyses were restricted to those patients who had similar response–character styles on both the MMPI and the Rorschach. For instance, the average correlation between the Rorschach *DEPI* and five MMPI–2 scales of dysphoric affect was $r = .58$ when only those patients with similar response–character styles were considered. Further, there were strong negative heteromethod validity coefficients for scales of emotional distress ($r = -.55$) when the analyses were restricted to patients who had opposing response–character styles on each method. Although the latter findings may seem counterintuitive, such testing patterns correspond to expected dynamic operations and certain diagnostic conditions (Finn, 1996; Meyer, 1997b).

Meyer (1997b) identified response–character styles using the first unrotated principal component from the MMPI and from the Rorschach. While this approach follows a long tradition in research on method variance (see Campbell & Fiske, 1959; Edwards & Edwards, 1991), it also has at least three disadvantages. First, the Rorschach factor scale is somewhat challenging to calculate. Second, MMPI profiles and Rorschach structural summaries are generally not viewed in terms of their underlying factor structures. Third, and perhaps most important, the first factor from each method is defined by variables that are also considered to be markers of other constructs. For instance, the first factor of the Rorschach is partially defined by Color-Shading Blends, *FY, FC′, S, FV,* and *Morbid* responses, all of which are traditionally interpreted as indications of negative affect. While this is entirely as it should be from a factor-analytic perspective, it has raised the question of whether Meyer's analyses of convergent validity for dysphoria scales were confounded by statistical and conceptual artifacts (see Ganellen, 1996). Given the foregoing, it would be optimal to have adequate criteria for defining these styles that (a) are simple to calculate and implement, (b) rely on more commonly recognized MMPI and Rorschach indicators of response–character style, and (c) are not potentially confounded by alternative constructs. Devising and testing such criteria were the primary goals for this study.

Prior research also investigated whether methodological artifacts that result from selecting extreme groups for analysis (i.e., patients matched on response–character styles) may have artificially produced the convergent validity results (Meyer, 1997b). By applying a correction formula to the observed data, the analyses suggested methodological artifacts could not account for the convergent findings. Still, methodological confounds remain critical issues to consider in this line of research. Thus, two new approaches were used to test the integrity of the results. First, it was determined if Rorschach and MMPI variable pairs that should be uncorrelated actually remained uncorrelated after patients were selected for their stylistic patterns. If conceptually meaningless variable pairs produced correlations of the same magnitude as conceptually meaningful variable pairs, then the "convergent validity" analyses would document nothing of substance.

Second, the general hypothesis guiding this research is that Rorschach and MMPI constructs will be correlated to the extent that test interaction styles are aligned across methods. To date, this proposition has been tested by selecting groups of patients who either have similar or dissimilar styles on each method. However, the same principal could be demonstrated without selecting particular groups of patients if one had access to enough patient samples. The trick would be to ensure the samples naturally varied in the extent to which Rorschach and MMPI first factors were correlated. If some samples had strong positive correlations, some had near zero correlations, and some had strong negative correlations, then one could test the hypothesis that the Rorschach and MMPI constructs would be differentially correlated to a similar extent across these samples. Even without having access to many independent samples of patients, this hypothesis could still be tested by repeatedly sampling from a single large population, which is the approach used in this study.

There was also another reason for initiating this study. Recently, Krishnamurthy, Archer, and House (1996) were unable to replicate some of the adult findings on MMPI–Rorschach convergent validity in a sample of adolescent patients. They examined validity as a function of both sets of criteria used in the adult literature: (a) response–character styles affecting both methods and (b) Rorschach *R*. In contrast to the adult research, Krishnamurthy et al. found "very limited associations between conceptually related MMPI–A and Rorschach variables" (p. 179).

In terms of broad response–character styles, Krishnamurthy et al. (1996) only examined patients with similar styles on both methods. Unfortunately, they did not report specific results for the variables used in the adult research. Instead, they calculated 237 Rorschach–MMPI correlations and found only 6.3% were statistically significant. This level was only marginally above the 1.0% expected by chance (given their $\alpha = .01$), and it suggested the adult findings did not replicate.

Krishnamurthy et al.'s (1996) results raise several possibilities. First, the adult research may have reflected chance findings that will not replicate in other samples—adult or adolescent. Second, personality organization may be sufficiently different in adults and adolescents that the findings only hold for adults. Third, the criteria used to define response–character styles with adolescents may have been inadequate to generate convergent validity. Krishnamurthy et al. did not employ the factor criteria used previously, presuming it would have been inappropriate to apply a factor structure derived from adults in an adolescent sample. Instead, they devised different but "conceptually similar" (p. 189) criteria. If the two studies produced very different results while using roughly equivalent procedures, then one should be quite cautious about accepting the general notions that have been proposed regarding cross-method convergence (Meyer, 1996, 1997b). However, if the adolescent criteria selected very different patients, then the adult findings may still be valid pending genuine replication. To evaluate this prospect, this study examines the similarity between the adult and adolescent criteria. Regardless of similarity, this study will also determine whether the adolescent criteria could produce convergent validity in Meyer's (1997b) sample of adult patients. If the criteria failed to work in the original adult sample, there would be no reason to expect them to work with adolescents.

In terms of convergent validity based on *R*, Krishnamurthy et al. (1996) found no significant correlations between the *DEPI* and MMPI–A depression scales in brief or lengthy records. The criteria used to select short and long records were quite similar to those in the adult research, so this did not appear to be a confound. Given that the initial results with *R* (Meyer, 1993) were unexpected and found with relatively small samples ($ns \approx 30$), the final purpose of this study was to conduct an analysis of the moderating influence of *R* in a larger sample of adults.

METHOD

Participants and Measures

This study employed the same participants and procedures reported by Meyer (1997b), so only brief descriptions will be provided here. The sample consisted of 362 patients who had completed an MMPI–2 and Rorschach. A total of 52% were psychiatric inpatients, 30% were psychiatric outpatients, 15% were general medical patients, and a few patients were drawn from other settings (3%). The average age was 34.9 years ($SD = 11.5$, range = 17–72 years); 55% were women; 60% were White, 32% African American, 4% Asian, and 4% other.

All Rorschachs were administered and scored according to Comprehensive System guidelines (Exner, 1993). Reliability was calculated with two independent coders using 63 protocols. Combining the agreement rates reported earlier with the base rates for each score in this sample (see Meyer, 1997a), kappa values were calculated to correct for chance agreement. Estimated kappa values were "good" or "excellent" (see Fleiss, 1981), ranging from a low of .63 for cognitive special scores to a high of .91 for location/space and popular ($M = .79$). All MMPI–2s were administered at roughly the same time as the Rorschach and computer scored.

Conceptually meaningful variables pairs. For the analyses related to broad response–character styles, dependent variables were considered in three content clusters: affective distress, psychotic processes, and interpersonal suspiciousness. All variables in this study matched those used in the prior adult research. Rorschach measures of emotional distress included the *DEPI* and the Suicide Constellation (*S–CON*). MMPI–2 variables included Scales *2* and *7,* Depression (*DEP*), Anxiety (*ANX*), and the Negative Emotionality/Neuroticism scale from the Personality Psychopathology Five (PSY–5–*Neg;* Harkness, McNulty, & Ben-Porath, 1995). The Rorschach measure of psychotic processes was the Schizophrenia Index (*SCZI*), while MMPI–2 measures included Scale *8, BIZ,* and the Psychoticism scale from the PSY–5 (PSY–5–*Psy;* Harkness et al., 1995). The Rorschach measure of interpersonal wariness was the Hypervigilance Index (HVI) and the MMPI–2 scales included Scale *6,* Cynicism (*CYN*), Social Discomfort (*SOD*), and the Inability to Disclose component of the Negative Treatment Indicators Scale (*TRT2;* Ben-Porath & Sherwood, 1993). In addition to examining each of these variables individually, all of the scales targeting a common construct were transformed to *z* scores and aggregated to form more reliable and valid composite measures (Tsujimoto, Hamilton, & Berger, 1990). Thus, the two Rorschach measures of negative affect were aggregated, as were the various MMPI scales targeting each of the three general constructs.

The analyses related to *R* focused on the construct of affective distress, because it displayed the strongest convergent validity in the previous study (Meyer, 1993). To parallel the prior research exactly, this analysis used the Rorschach *DEPI* and the following MMPI–2 scales: Scale *2,* Scale *0, DEP, SOD,* Low Self-Esteem (*LSE*), Subjective Depression (*D1*), and Work/Interference (*WRK*).

These analyses only examined the variables listed. Findings were not culled from a larger array of results that may have led to different conclusions.

Conceptually unrelated variable pairs. To evaluate the discriminant validity of the processes being considered here, I selected a range of variables that bore little conceptual relation to each other. Initially, a limited set of Rorschach–MMPI variable pairs was identified. However, these analyses were abandoned after I realized all variables did not have strong positive correlations with their respective first factors. Subsequently, new variables were selected, all of which correlated positively with either the Rorschach or MMPI first factor. Rather than specifying a large number of unique variable pairs, sets of Rorschach and MMPI variables were designated. In each set, I chose variables that had approximately normal distributions and had not been used in the convergent analyses. The variables in both sets were expected to be globally independent across all comparisons.

Thirteen Rorschach scores were selected: whole locations (*W*), detail locations (*D*), ordinary developmental quality (*DQo*), ordinary form quality (*FQo*), animal movement (*FM*), Pairs, the sum of all human content, whole human content (*H*), whole animal content (*A*), idiographic content (*Id*), passive movement (*p*), personalized responses (*PER*), and cooperative movement (*COP*). These variables had an average correlation of .44 with the Rorschach first factor, which was similar to the average correlation found with the four Rorschach variables used in the meaningful convergent validity analyses ($r = .48$). Eleven MMPI variables were also selected: Scale *1*, Scale *3*, Somatic Complaints (*Hy4*), Ego Inflation (*Ma4*), Generalized Fearfulness (*FRS1*), Health Concerns (*HEA*), Gastrointestinal Symptoms (*HEA1*), Explosive Behavior (*ANG1*), Addiction Admission (*AAS*), Antisocial Behavior (*ASP2*), and Aggression (Psy–5–*Agg*). These variables had an average correlation of .52 with the MMPI first factor. Although this was lower than the average first factor correlation for the 12 variables used in the meaningful convergent analyses ($r = .78$), it was not possible to select MMPI scores with substantially larger first factor correlations. The 13 Rorschach scores and 11 MMPI scores produced 143 variable pairs, which were labeled *Conceptually Unrelated Variable Pairs—Selected for High First Factor Correlations* (CUVP–*HighFF*). As expected, these variables were essentially uncorrelated in the full sample of 362 patients (mean $r = -.03$; range = –.13 to .11).

Next, to provide a second set of meaningless variables for comparison, I returned to the variables initially considered—those selected without regard to first factor loadings. From these, five Rorschach and five MMPI variables were chosen to replace ones from the previous sets. The 13 Rorschach scores now consisted of *W, D,* pure form (*F*), *DQo, FQo, FM, Pairs,* the adjusted *D* score (*AdjD*), popular, *A, Id,* the affective ratio (*Afr*), *PER,* and good form quality percent (*X+%*). These variables had an average correlation of .27 with the Rorschach first factor. The 11 MMPI variables were now: Scale *3*, Need for Affection (*Hy2*), *Ma4,* Gender Masculinity (*GM*), Gender Femininity (*GF*), Social Responsibility (*Re*), *FRS1, HEA1, ASP2,* Psy–5–*Agg,* and Positive Emotionality (Psy–5–*Pos*). These variables had an average correlation of –.02 with the MMPI first factor (range = –.76 to .65). As before, these scores produced 143 variable pairs that were essentially uncorrelated in the full sample of patients (mean $r = -.01$; range = –.19 to .15). These variables were labeled *Conceptually Unrelated Variable Pairs—Not Selected for First Factor Correlations* (CUVP–*NotFF*).

Procedures

Defining broad response–character styles with factor-based criteria. The procedures for defining response–character styles using the first principal components from the Rorschach and the MMPI–2 were discussed in detail by Meyer (1997b). A summary of these criteria can be found in the first column of Table 2. Briefly, because Welsh's Anxiety Scale (*A*) was designed to quantify the first MMPI factor, it was used in these analyses instead of generating factor scores from the present sample. Nonetheless, *A* was found to be an excellent marker of the

TABLE 2
Three Criteria Sets Used to Define Broad Response–Character Styles on the MMPI–2 and Rorschach

	Factor Criteria	*Profile Criteria*	*Krishnamurthy et al. (1996) Criteria*
Dilated response style			
MMPI–2	Welsh's Anxiety (raw) > 21	*K T* score < 50 and *F T* score > 58	*F T* score > 64
Rorschach	Highest 1/3 on the Response Engagement Factor Scale	*R* > 21 and *Lambda* < 0.55	*Lambda* < 1.08
Constricted response style			
MMPI–2	Welsh's Anxiety (raw) < 11	*K T* score > 50 and *F T* score < 58	*L* or *K T* score >55
Rorschach	Lowest 1/3 on the Response Engagement Factor Scale	*R* < 21 and *Lambda* > 0.55	*Lambda* > 1.0
Patients in this sample having similar MMPI and Rorschach response styles (%)	24.0	14.9	40.6
Patients in this sample having opposing MMPI and Rorschach response styles (%)	21.5	14.1	51.7

MMPI–2's first unrotated principal component in this sample, with a loading of .95 on this dimension. The formula for calculating factor scores for the Rorschach's Response–Engagement scale was initially derived from a large sample of college students (Meyer, 1992b). In the current psychiatric sample, this variable was the best marker of the Rorschach's first principal component, having a loading of .96. Conceptually, one end of the first principal component from each method is characterized by defensive withdrawal, cognitive-emotional simplicity, or denial (see Table 1). The other pole is characterized by excessive engagement, heightened sensitivity, or overreporting of problems. For the sake of simplicity, these poles are termed *constricted* and *dilated,* respectively.

Defining broad response–character styles with profile scores. In factor analyses of the Rorschach (Meyer, 1992a, 1992b), *R* is consistently a key variable for defining the dilated end of the first principal component. Conversely, *Lambda* or *Pure F%* consistently has a strong loading on the constricted end of this dimension. Because *R* is a strong positive marker of the first factor and *Lambda* is a decent negative marker, both were selected to define this dimension. Exner (1993) indicated that these two variables can be considered markers of response–character styles and these variables also quantify the pervasive response–character style dimension Rappaport, Gill, and Schafer (1968) believed was essential to consider before interpreting other Rorschach scales. Several studies relate these variables to task engagement or behavioral defensiveness and constriction (Adams & Cooper, 1961; Belmont & Birch, 1962; Epstein, Lundborg, & Kaplan, 1962; Finch, Imm, & Belter, 1990; Jourard, 1961; Masling, 1960; Shatin, 1958; Wohl, 1957).

To evaluate the adequacy of *R* and *Lambda* as markers of the Rorschach's first factor, I conducted a principal components analysis of nonredundant scores for location, space, developmental quality, determinants, special scores, and interpretive ratios ($N = 442$). Consistent with expectations, *R* had a strong loading (.70) on the first unrotated principal component, while *Lambda* had the largest negative loading (–.40).[4]

[4]*Lambda* often has a skewed and kurtotic distribution in clinical samples because it is calculated as a proportion. When *Pure F* is ≤ 50% of *R,* this proportion has a fixed range between zero and 1.0. However, it has an unlimited upper range constrained only by *R* itself. As such, just one patient with a large proportion of pure form responses can skew the *Lambda* distribution severely. For instance, a patient with 17 out of 18 *Pure F* responses will produce a *Lambda* value of 17.0, which is a dramatic outlier. This problem is easily fixed by calculating the percentage of responses that are pure form (i.e., *Pure F%*). This variable always has a more normally shaped distribution than *Lambda.* Consequently, *Pure F%* is actually a better marker for the negative pole of the Rorschach's first factor. It was not used in these analyses because the goal was to develop marker variables that could be obtained directly from the structural summary.

There is a much longer history of research on the factor structure of the MMPI. The basic findings are not in dispute, because the first unrotated principal component of the MMPI is quite stable. On the dilated pole, this component is defined by Welsh's *A* scale as well as by *Sc, Pt,* and *F.* The constricted pole is most strongly defined by *K* (Butcher et al., 1989; Dahlstrom, Welsh, & Dahlstrom, 1975; Graham, 1993; Greene, 1991). Although this factor is quite robust, there has been extensive debate about its meaning. Some have argued that it reflects genuine personality characteristics related to anxiety, neuroticism, negative affectivity, or global psychopathology (e.g., Block, 1965; Johnson, Butcher, Null, & Johnson, 1984; Welsh, 1956). Others have argued that it is best viewed as an index of general response styles like acquiescence and social desirability (Edwards, 1957, 1964; Edwards & Edwards, 1991; Messick & Jackson, 1961). A simple resolution is to recognize that this dimension, and the first factor of the Rorschach, measure genuine trait characteristics for some people but also the bias and error of response–character styles for others.

Because *F* is a strong positive marker of the MMPI's first factor while *K* is the strongest negative marker, both variables were selected to define this dimension. Importantly, every major text on the MMPI identifies *F* and *K* as prominent indicators of response style (e.g., Archer, 1992; Butcher & Williams, 1992; Graham, 1993; Greene, 1991). To evaluate the utility of these marker variables, the basic scale intercorrelations from the MMPI–2 normative sample (Butcher et al., 1989, Table F–1) were used to generate a factor analysis. For males, *F* had a loading of .73 on the first unrotated principal component (third highest after *Sc* and *Pt*), whereas *K* had the strongest negative loading on this factor (–.69). The data were almost identical for females. The loading for *F* was .74 (third highest), while *K* had the largest negative loading (–.63). Thus, these two scales are relatively good markers for the first unrotated principal component of the MMPI.

To classify patients as dilated or constricted, I employed the median values found in this sample as cut points.[5] These values can be identified by examining the second column of Table 2. For instance, the median *T* score for the *K* scale in this sample was 50. Consequently, values less than 50 were taken as one indication that the patient was dilated, whereas values greater than 50 were taken as one indication that the patient was constricted.

Two aspects of this procedure should be noted. First, the median values for *K, R,* and *Lambda* (50, 21, and .55, respectively) corresponded quite closely to the median values observed in nonpatient normative samples (50, 23, and .56, respectively; see Butcher et al., 1989; Exner, 1993). Because these variables reflect response–character styles that should be independent of general psychopathology, these are anticipated results. The finding was different for the *F* scale, however. In

[5]As with *Lambda,* the distribution for *F* is generally skewed in clinical samples. Consequently, for both variables the median is a better statistic for describing central tendency than the mean.

this sample the median *T* score was 58, which was elevated .8 standard deviations above the normative nonpatient mean of 50. This is also to be expected because *F* is simultaneously a measure of psychopathology and of response–character styles (e.g., Archer, 1992; Graham, 1993; Greene, 1991). Theoretically, psychiatric patients should be elevated on *F* because of genuine psychopathology. Had the nonpatient median been used to define styles, a greater proportion of trait variance would have been erroneously treated as if it were stylistic method variance. Overall, the rationale presented here is that deviations from nonpatient normative values can be used to identify response–character styles for *K, R,* and *Lambda.* However, for the *F* scale, styles should be identified after taking into account the variance on *F* that reasonably could be attributed to genuine psychopathology. For a psychiatric sample, the cut point should always be above the value found in the normative population.

Second, standard clinical texts have defined response styles using criteria such as *T* scores for *K* greater than 65 (Butcher et al., 1989), *R* values less than 14, or *Lambda* values greater than .99 (Exner, 1993). From such a perspective, the median splits used here may seem overly liberal criteria for defining response–character styles. However, they actually are not. The reasoning is twofold. First, the goal is not to identify test protocols that are clinically suspect or "invalid." Rather, the goal is to estimate the location of patients on the first unrotated principal component in order to identify their broad stylistic approach to completing each test. Second, no scale is used as an independent measure of response–character styles. Rather, two scales from each method are combined. The proportion of people identified by the combination of these scales is much less than 50% of the sample. As can be seen from the bottom rows in Table 2, the profile criteria actually identify a smaller proportion of patients than the factor criteria employed earlier.[6]

Defining broad response–character styles with Krishnamurthy et al.'s (1996) criteria. The last column in Table 2 lists the criteria used by Krishnamurthy et al. (1996) to define broad response–character styles. As can be seen, these authors generally used a single variable for this purpose, whereas the profile scores use one from each pole of the first principal component. Also, even when the same variables are used, quite different cut points identify the styles (e.g., dilated *Lambda* values of < .55 vs. < 1.08).

[6]One reviewer questioned the appropriateness of these criteria and recommended more stringent cutoffs to define response–character styles. It is certainly the case that the values used here are more liberal than the cut points recommended for identifying invalid data in clinical practice. It is also the case that the criteria recommended here may ultimately prove to be suboptimal. Additional research using alternative criteria may help to refine the best procedures for measuring each test's first principal component using *F* and *K* and *R* and *Lambda* (or *Pure F%*).

Determining the similarity of response–character styles across methods. For subsequent analyses, patients were classified as having similar response styles on the MMPI and Rorschach if they met the dilated criteria for each method or if they met the constricted criteria for each method. Patients were classified as having opposing response–character styles if they met dilated criteria for one test and constricted criteria for the other. These designations were made independently for each of the three criteria sets listed in Table 2.

Defining Rorschachs with simple response frequency. Meyer (1993) defined protocols as follows: brief, R = 14 to 17; average, R = 21 to 24; lengthy, $R >$ 28. In the initial study, these groups had *n*s of 22, 29, and 30, respectively. The sample from this study had *n*s of 90, 73, and 74, respectively (125 protocols had R = 19–20 or 25–28, so they were not classified). Krishnamurthy et al. (1996) defined R subgroups using two slightly different criteria sets. However, their criteria were highly correlated with Meyer's (r_S = .85 and .91) and should have produced similar findings.

Repeated sampling from the fixed population. It is well known that individuals samples—particularly those that are relatively small (e.g., n = 30)—can produce fluctuating or imprecise results. This is a fundamental principle behind inferential statistics, and it is a natural consequence of the random processes that are inherent in drawing samples from populations (see Hays, 1981). For the present purposes, these sample-to-sample fluctuations were harnessed to evaluate the more general relation between response–character styles and the convergence of Rorschach and MMPI constructs. By repeatedly drawing samples from a common population it was expected that the association between the Rorschach and MMPI first factors would fluctuate from one sample to the next. This natural variation across samples provides a way to test the hypothesis that correlations between Rorschach and MMPI constructs fluctuate across samples in tandem with the degree of association between response–character styles.

To determine this, 300 patient samples were created by taking 300 random draws from the full population of 362 patients. On each draw, two correlations were computed. Variable A consisted of the correlation between the Rorschach and MMPI first factors. Variable B was created by correlating each of the 17 conceptually meaningful Rorschach and MMPI construct pairs and then averaging the 17 correlations to obtain a summary index of convergent validity among constructs. The correlation between Variable A and Variable B was then calculated across the 300 subsamples, treating each subsample as a single observation or "subject."

These analyses were conducted twice. On the first occasion, samples of 30 patients were randomly selected on each of the 300 draws. These samples were roughly the same size as those used by Meyer (1993) and were small enough to ensure substantial variability in the extent to which the first factors were (i.e., appeared to be) correlated in the 300 samples. On the second occasion, samples of 75 patients were randomly selected on each of the 300 draws. Although of moderate size, samples this large should show less random variation in first factor correlations.

To appreciate these analyses, it is essential to keep in mind the relevant population parameters. In this case, the population consisted of all 362 patients with MMPI and Rorschach data. As reported by Meyer (1997b), Variable A, the correlation between Rorschach and MMPI first factors, had a value of $r = .04$ in the population, whereas Variable B, the average correlation among the 17 construct pairs, had a value of $r = .03$. Thus, theoretically, for every sample obtained on one of the 300 random draws, Variable A should have a value of .04 and Variable B should have a value of .03. This does not happen, however, because sampling error ensures that random factors affect each draw, such that the values for Variable A and B on each draw only approximate the population parameters. Smaller samples produce more sampling error than larger samples, so 300 random draws of 30 patients at a time will produce more widely disparate values for Variable A and Variable B than will 300 random draws of 75 patients.

RESULTS AND DISCUSSION

Broad Response–Character Styles Defined by Factor and Profile Scores

Associations among criteria. To quantify the association between the three criteria sets proposed for identifying response–character styles (Table 2), patients were classified as dilated, indeterminate (i.e., without a clear response–character style), or constricted.[7] The chance-corrected (κ) rates of classification agreement are given in Table 3. Considering first the MMPI criteria, the profile scores had a good degree of association with the factor criteria (see Fleiss, 1981). However, the association between each of these criteria sets and Krishnamurthy et al.'s (1996) were only fair. With respect to the Rorschach, the association between the factor and profile criteria was fair at best. The association between each of these classification criteria and those used by

[7]For this analysis, the Rorschach criteria used by Krishnamurthy, Archer, and House (1996) were modified slightly to make them mutually exclusive categories. *Lambda* < 1.08 was considered dilated, *Lambda* > 1.08 (not 1.0) was considered constricted, and *Lambda* = 1.08 was considered indeterminate.

TABLE 3
Kappa Coefficients Indicating Classification Agreement Among the Three Criteria Sets Used to Define Broad Response–Character Styles on the MMPI–2 and the Rorschach

Scale and Criteria	*Factor Criteria*	*Profile Criteria*
MMPI–2		
Profile criteria	.61*	
Krishnamurthy et al. (1996) criteria	.40*	.46*
Rorschach		
Profile criteria	.37*	
Krishnamurthy et al. (1996) criteria	.22*	.11*
Similar styles on both methods		
Profile criteria	.45*	
Krishnamurthy et al. (1996) criteria	.18*	.14*
Discordant styles on both methods		
Profile criteria	.30*	
Krishnamurthy et al. (1996) criteria	.03	.02
Overall classification on both methods		
Profile criteria	.31*	
Krishnamurthy et al. (1996) criteria	.11*	.10*

Note. $N = 352$.
*$p < .001$.

Krishnamurthy et al. were poor. Considering the groups of central concern to the subsequent analyses, when patients identified as having similar styles were compared to the remaining patients, classification agreement between the factor and profile criteria was fair. However, the agreement between these two criteria sets and the adolescent criteria was poor, producing kappa values of .18 and .14, respectively. When patients identified as having discordant styles were compared to the remaining patients, agreement between the factor and profile criteria was poor to fair. The agreement between these two criteria sets and the adolescent criteria was quite poor, producing kappa values near zero. When all classifications were considered simultaneously, the agreement between factor and profile criteria was poor to fair. The agreement between the adult and adolescent criteria sets was again quite poor.

These analyses demonstrated two things. First, the profile criteria are not synonymous with the factor-based criteria used in prior adult research. Substantially different groups of patients are identified with these two criteria sets, particularly for discordant styles. Consequently, when the profile criteria are used to select patients for an analysis of Rorschach and MMPI convergent validity, the results will provide a relatively independent check on the hypothesis because convergence will be tested with substantially different samples of patients. Second, the factor and profile criteria are both very different than the criteria proposed by Krishnamurthy et al. (1996). In fact, when considering who actually

TABLE 4
MMPI–2 and Rorschach Convergent Validity for Three Constructs Using the Three Criteria Sets to Define Patients With Similar Response–Character Styles

	Response–Character Styles Defined By					
	Factor Criteria[a]		*Profile Criteria*[b]		*Krishnamurthy et al. (1996) Criteria*[c]	
Negative Affect	*DEPI*	*S–CON*	*DEPI*	*S–CON*	*DEPI*	*S–CON*
Scale 2	.42***	.50***	.19	.36**	–.09	.04
Scale *7*	.62***	.67***	.46***	.51***	.11	.12
DEP	.66***	.63***	.46***	.48***	.10	.12
ANX	.61***	.62***	.46***	.51***	.10	.11
PSY–5–*Neg*	.59***	.62***	.47***	.47***	.09	.09
Psychosis	*SCZI*		*SCZI*		*SCZI*	
Scale *8*	.54***		.46***		.19*	
Bizarre Mentation	.46***		.44***		.11	
PSY–5–Psychoticism	.48***		.46***		.11	
Interpersonal Wariness	*HVI*		*HVI*		*HVI*	
Scale *6*	.45***		.37**		.15	
Cynicism	.22*		.27		–.04	
Social Discomfort	.38***		.40**		.00	
Inability to Disclose	.44***		.49***		.09	

Note. DEPI = Depression Index; S–CON = Suicide Constellation; *DEP* = Depression Content Scale; *ANX* = Anxiety Content Scale; PSY–5–*Neg* = Personality Psychopathology Five Negative Emotionality/Neuroticism Scale; SCZI = Schizophrenia Index; HVI = Hypervigilance Index.
[a]$n = 87$. [b]$n = 54$. [c]$n = 147$.
*$p < .05$. **$p < .01$. ***$p < .001$.

becomes classified as having similar approaches to completing the MMPI and Rorschach, there is only minimal overlap between the adult and adolescent criteria. As a result, the adolescent criteria may not generate the same pattern of convergent validity.

MMPI and Rorschach convergent validity. The next set of analyses determined the effectiveness of each criteria set for producing convergent validity by selecting patients with similar response styles on both methods and correlating Rorschach and MMPI–2 scales. Table 4 presents results for the three target constructs.[8] It is readily apparent that the factor and profile criteria generated

[8]The factor results were reported by Meyer (1997b) and are presented here for comparison.

convergent validity between the MMPI–2 and Rorschach, whereas the adolescent criteria did not. Across all constructs, the factor criteria generated an average validity coefficient of $r = .52$. The profile scores were somewhat less effective, although they still generated strong convergent validity, with an average correlation across constructs of $r = .43$. The criteria employed in the adolescent research were ineffective for producing cross-method convergent validity in this adult sample, having a mean coefficient across constructs of $r = .08$.

The top section of Table 5 presents summary information related to the preceding analyses. In addition to presenting the average coefficients for each construct, the table also reports findings for the composite variables. The latter readily demonstrates how aggregation reduces error and produces variables with validity coefficients that are higher than the average of the individual coefficients (e.g., factor criteria for negative affect, $r = .71$ vs .59).

TABLE 5
Summary Coefficients (Means and Composites) for Construct Pairs Hypothesized to be Correlated and for the Mean of Two Sets of 143 Construct Pairs Hypothesized to be Uncorrelated

	Criteria Used for Selection					
	Factor		*Profile*		*Krishnamurthy et al. (1996)*	
Style and Variables	*M*	*Composite*	*M*	*Composite*	*M*	*Composite*
Similar styles[a]						
Negative Affect	.59***	.71***	.44***	.50***	.08	.19*
Psychosis	.49***	.53***	.45***	.49***	.14	.20*
Wariness	.37***	.50***	.38**	.50***	.05	.11
CUVP-*NotFF*	–.03		.02		.00	
CUVP-*HighFF*	.03		.18		.00	
Discordant styles[b]						
Negative Affect	–.54***	–.62***	–.26	–.31*	–.04	–.03
Psychosis	–.19	–.20	–.16	–.18	.06	–.03
Wariness	–.28*	–.34**	–.24	–.32*	.09	.12
CUVP-*NotFF*	.01		.02		.00	
CUVP-*HighFF*	–.18		–.25		–.03	

Note. CUVP = conceptually unrelated variable pairs; *NotFF* = not selected for first factor correlations; *HighFF* = selected for high first factor correlations. Mean coefficients are based on 10 construct pairs for Negative Affect, 3 construct pairs for Psychosis, 4 construct pairs for Interpersonal Wariness, and 143 construct pairs for each set of CUVP.

[a]*ns* = 87, 54, and 147 for the factor, profile, and Krishnamurthy et al. (1996) criteria, respectively. [b]*ns* = 78, 51, and 187, respectively.

*$p < .05$. **$p < .01$. ***$p < .001$.

For those patients classified as having opposing response–character styles, the lower section of Table 5 indicates that the factor and profile criteria produced the expected negative correlations among the meaningful constructs, though the adolescent criteria did not. However, even with the factor and profile criteria, validity coefficients for the construct of psychosis were smaller ($r = -.20$ and $-.18$) and not statistically significant. Furthermore, the profile criteria produced only medium sized coefficients (Cohen, 1988) for negative affect (–.31) and wariness (–.32) and were just marginally significant at a two-tailed probability level. The inferiority of the profile criteria relative to the factor criteria is most likely due to their small correspondence (i.e., $\kappa = .30$; see Table 3).

Table 5 also presents results for the two sets of 143 variable pairs expected to remain uncorrelated after selecting patients on response–character styles. As anticipated, for each of the six groups presented in Table 5 (i.e., factor criteria for similar styles, factor criteria for discordant styles, etc.), the average coefficients for the 143 variable pairs in each set were not statistically significant. However, a more important consideration is whether meaningful variable pairs (e.g., those targeting negative affect) produced larger[9] correlations than those obtained from the sets of CUVPs.

For similar styles, the factor and profile criteria produced convergent correlations substantially larger than those produced by either set of CUVP. All the meaningful constructs had $r > .48$, whereas all the CUVPs had $r < .19$. Thus, when either the factor or profile criteria are used to identify similar test interaction styles across methods, there should be substantial Rorschach–MMPI construct convergence. This construct convergence does not appear to be caused by the simple alignment of methodological artifacts because neither set of 143 CUVPs produced similarly large correlations.

For the discordant test interaction styles, the same global conclusions do not hold. Although both the factor and profile criteria produced convergent coeffi-

[9]There are sound procedures for determining if nonoverlapping correlations obtained from the same participants are significantly different from each other (see Raghunathan, Rosenthal, & Rubin, 1996; Steiger, 1980). However, procedures have not been worked out for comparing the average of nonoverlapping sets of correlated correlations (i.e., comparing the average of the negative affect coefficients to the average of the coefficients expected to be uncorrelated). Steiger (1980) and Raghunathan, Rosenthal, and Rubin (1996) presented approaches that can be used for contrasting individual variables. However, these procedures require all possible variable pairs to be computed and transformed to determine the statistical significance of the differences. If the three meaningful sets of constructs were to be compared with the CUVP coefficients, there would be a total of 40 variables under consideration. This would require a matrix containing 780 off-diagonal correlations for use in the subsequent transformations to determine significance. Furthermore, a separate matrix would have to be computed for each of the six analyses reported in Table 5 (i.e., similar and discordant styles using each criteria set). Finally, separate large scale analyses would be required to determine if the composite correlations differed from the average of the CUVP coefficients. Given the forgoing, the statistical significance of differences between average correlations was not determined.

cients consistently larger than the CUVP–*NotFF* set, the same differentiation was not always present when comparing the conceptually matched constructs to the CUVP–*HighFF* set. Consistent with findings reported earlier for the factor criteria (Meyer, 1997b), the construct of negative affect demonstrated meaningfully larger convergent validity coefficients ($r = -.62$) than what be attributed to methodological artifacts ($r = -.18$ for CUVP–*HighFF*). In this analysis, the construct of wariness also seemed to have some capacity to be incrementally differentiated from these meaningless variable pairs ($r = -.34$ vs. $-.18$). However, this should be viewed tentatively, because the difference is relatively small and because an alternative way to correct methodological artifacts did not support this conclusion (Meyer, 1997b, p. 322).

Unlike the factor criteria, the new profile criteria did not produce coefficients for any of the meaningful construct pairs ($rs = -.31, -.18, -.32$) that were clearly larger than the average correlation from the CUVP–*HighFF* set ($r = -.25$). Thus, when the profile criteria were used to identify a constricted style on one method and a dilated style on the other, the results were undifferentiated. Negative correlations were present among all variable pairs with strong first factor loadings, regardless of their conceptual alignment. Similarly, when using the factor criteria, Rorschach and MMPI scales measuring psychosis ($r = -.20$) did not produce coefficients that could be differentiated from meaningless variable pairs when all of the latter also had strong first factor correlations ($r = -.18$).

Revisiting MMPI and Rorschach Convergent Validity Based on Rorschach Response Frequency

DEPI and MMPI–2 scales were correlated within each of the subgroups defined by *R*. Brief records had an average convergent correlation of $r = -.09$. This is in contrast to the mean $r = -.25$ found in the initial research (Meyer, 1993).[10] The average length protocols now produced a mean correlation of $r = -.03$, rather than the value $r = -.21$ found in the smaller sample. Lengthy records now produced an average $r = .11$, rather than the value of $r = .34$ found earlier.

Based on these findings it appears the initial results (Meyer, 1993) regarding *R* as a moderator of Rorschach and MMPI convergent validity were artifactual. The findings did not replicate when Krishnamurthy et al. (1996) used very similar criteria in an adolescent sample, and they do not replicate in an expanded sample of adults drawn from the same population as the initial study—and that even included

[10] In the process of calculating these results, I discovered there had been minor data entry errors in the original research (Meyer, 1993). Although these errors did not alter the substance of the original research, please see Meyer (1998) for a more complete discussion of the discrepancies.

some patients from the previous analysis. Thus, when considered in isolation, the variable *R* does not govern cross-method convergent validity.

Exploring the failure of the simple response frequency criteria. Given these findings it is important to understand why variations in *R* produced convergent validity coefficients in the initial study but not in the expanded sample. More recent theorizing suggests the answer should be found by examining response–character styles (Meyer, 1997b). Specifically, if MMPI–2 and Rorschach stylistic variables moderate convergent validity, then a sample of patients matched on response–character styles should produce positive convergent validity coefficients among trait constructs. Conversely, a sample of patients that happens to have opposing response–character styles across methods should generate negative convergent validity coefficients.

Thus, if the hypothesis is generally correct, the initial group of low *R* protocols should have had a stronger negative correlation between the first unrotated principal components from the MMPI–2 and Rorschach than the expanded group of low *R* protocols. This is because the depression scales had an average correlation of $r = -.25$ in the initial sample, but they had an average correlation of only $r = -.09$ in the expanded sample. By similar reasoning, there should have been a stronger positive correlation between the Rorschach and MMPI–2 first principal components in the initial sample of high *R* protocols than in the expanded sample of high *R* protocols. This is because the average validity coefficient in the initial analysis was $r = .34$, whereas it was only $r = .11$ in this analysis.

Results supported these hypotheses. The MMPI and Rorschach first factors had a correlation of –.46 in the initial low *R* group and of –.08 in the expanded sample. The two principal components had a correlation of .26 in the initial high *R* group and .16 in the expanded sample. Finally, the first principal components had a correlation of –.22 in the initial group of average *R* protocols and –.06 in the expanded sample. This is also as it should be, given the depression constructs had an average $r = -.21$ in the initial analyses and –.03 in the current sample.

To clarify these results further, the average degree of association among the Rorschach and MMPI depression constructs was correlated with the extent of association between the Rorschach and MMPI first principal components in each of the six subsamples (i.e., low *R*, average *R*, and high *R* groups in both the initial samples and the expanded samples). The observed correlation was .94 ($N = 6$, $p = .005$). This relation is graphically presented in Figure 1. Although the number of observations included in this analysis was quite limited, the results support the hypothesis that construct convergence is dependent on the alignment of response–character styles across methods. When Rorschach and MMPI response–character styles happen to be positively correlated, Rorschach and MMPI trait scales are positively correlated. When response–character

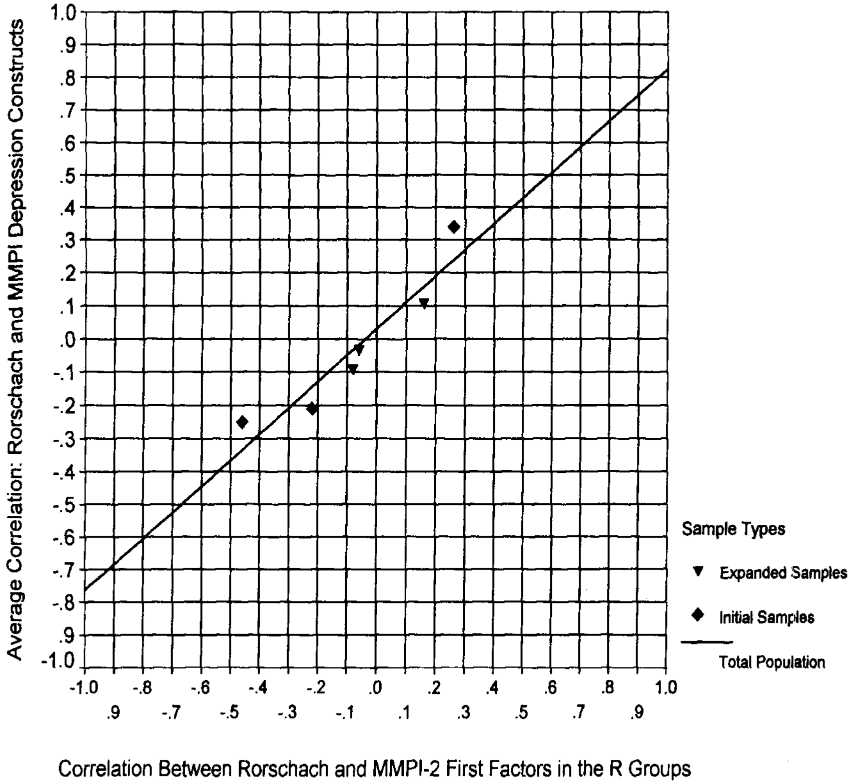

FIGURE 1 Rorschach and Minnesota Multiphasic Personality Inventory–2 (MMPI–2) construct convergence as a function of the correlation between the first unrotated principal components from the Rorschach and MMPI–2. Results from low-, average-, and high-*R* groups in two sets of analyses. Diamonds indicate the observations from the three initial *R* groups, and triangles indicate the observations from the three expanded *R* groups.

styles happen to be negatively correlated, Rorschach and MMPI trait scales are negatively correlated.

Thus, the hypothesis that broad response styles should foster Rorschach and MMPI–2 convergent validity can explain why the initial study found one pattern of relations while the expanded sample failed to find such a pattern. Initially, the subgroups defined by *R* coincidentally had a distinct pattern of correlations between MMPI and Rorschach response–character styles—the low *R* group just happened to have a fairly strong negative correlation between the Rorschach and MMPI first factors, whereas the high *R* group just happened to have a fairly strong positive correlation between the first factors. This caused the Rorschach and

MMPI construct pairs to be negatively correlated in the brief records and positively correlated in the subsample with lengthy records. However, in the expanded samples, the first factors tended to be uncorrelated in all groups. As a result, the Rorschach and MMPI construct pairs were also generally uncorrelated.

Response–Character Styles as General Moderators of Construct Convergence

Unlike Tables 4 and 5, which selected groups of patients based on prominent response–character styles, the preceding section indicates how these styles in general help determine the extent of association among conceptually similar Rorschach and MMPI constructs. Those data also indicate how individual samples—particularly those that are relatively small (e.g., $n = 30$)—can produce fluctuating or imprecise results.

The latter point leads to the final set of analyses, which capitalized on the random variation that emerges when taking samples from a population. As described earlier, sampling variation was explored as a way to test the basic hypothesis (i.e., response–character styles promote construct convergence) without deliberately selecting particular groups of patients. Instead, samples were randomly drawn from the full population 300 times and each time, two variables were computed: Variable A consisted of the correlation between Rorschach and MMPI first factors, and Variable B consisted of the average correlation among all the relevant Rorschach and MMPI construct pairs listed in Table 4. The correlation between Variable A and Variable B was then calculated across the 300 subsamples. This procedure was done twice, first with 300 random samples of 30 patients each and then with 300 random samples of 75 patients each.

When 30 patients were selected on each draw, the correlation between first factor correlations and construct correlations was .65 ($N = 300$, $p < .001$). When 75 patients were selected on each draw, the correlation between Variable A and Variable B was .66 ($N = 300$, $p < .001$).[11] These results clearly indicated that construct convergence was largely a function of response–character styles. The stronger the positive or negative correlation between first factors, the stronger the positive or negative correlation between Rorschach and MMPI trait constructs.

The relation between these variables is graphically presented in Figure 2 for random samples of $n = 30$, and Figure 3 for random samples of $n = 75$. In both fig-

[11]The results were no different when each correlation was transformed to Fisher's Z coefficient. For 75 patients, the correlation among the Z coefficients was .6646 rather than the value of .6612 with untransformed correlations; for 30 patients, the correlation was .6439 rather than .6503. Also, the results were very similar when the Rorschach *R–Engagement* scale and the MMPI–2 *A* scale were used rather than the first factor coefficients. For $n = 30$, $r = .63$; for $n = 75$, $r = .65$.

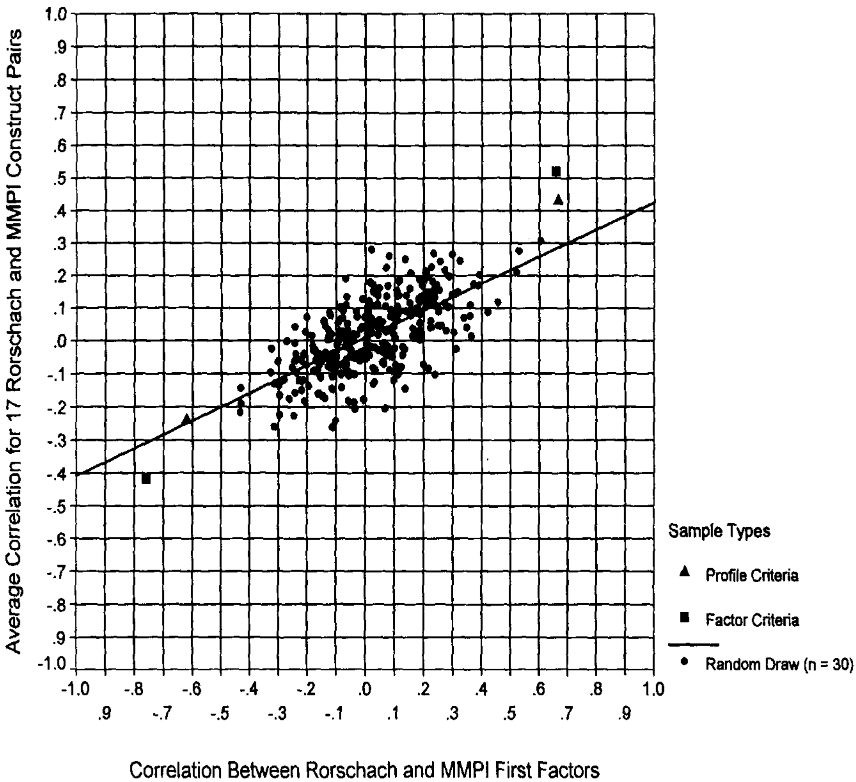

FIGURE 2 Demonstrating Rorschach and MMPI–2 construct convergence as a function of response–character styles using 300 randomly selected subsamples of 30 patients each. The samples selected on the basis of factor criteria (*ns* = 87 and 78) and profile criteria (*ns* = 51 and 54) are also shown.

ures, each of the 300 samples is represented by a circle. For reference purposes, the results for the deliberately selected samples are also reported. Samples selected by the factor criteria are indicated by squares, and those selected by the profile scores are indicated by triangles.[12]

In addition to providing strong support for the hypothesis that the convergence of Rorschach and MMPI constructs depends in part on response–character styles, several additional points should be noted about these data. First, Figure 2 indicates how small samples drawn from a single population generate very different results as one

[12]When these four data points were considered, the A–B correlation was .70 ($N = 304, p < .000$) in the $n = 30$ data set and .77 ($N = 304, p < .000$) in the $n = 75$ data set.

moves from sample to sample. Noticing how diffuse the cloud of circles in Figure 2 is relative to the those in Figure 3 vividly displays how sampling error affects small samples much more than large samples. In Figure 2, the randomly selected subsamples have first factor correlations that range from a low of –.43 to a high of .60—even though the true value in the population was actually $r = .04$! Similarly, for Variable B, measuring the average construct convergence, correlations ranged from a low of –.26 to a high of .31. This variability was evident even though the average correlation across the 17 construct pairs in the population was only .03. Although not shown in the figure, across the 300 random draws, the correlations for specific construct pairs also varied widely. Correlations ranged from a low of –.59 (*DEPI* with Scale *2*; *HVI* with *CYN*) to a high of .69 (*SCZI* with *BIZ*). Obviously, these values re-

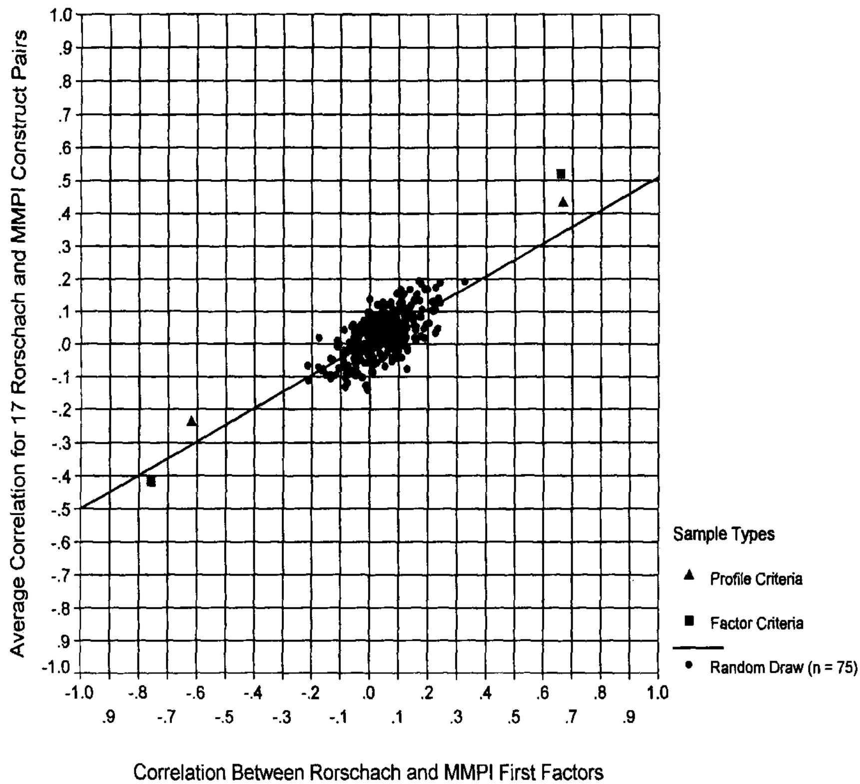

FIGURE 3 Demonstrating Rorschach and MMPI–2 construct convergence as a function of response–character styles using 300 randomly selected subsamples of 75 patients each. Also shown are the samples selected on the basis of factor criteria (*n*s = 87 and 78) and profile criteria (*n*s = 51 and 54).

flect a tremendous range of variability. Sample-to-sample variability is also evident in Figure 3, although the ranges are less extreme and cluster more tightly around the statistically expected values. First factor correlations varied from a low of –.22 to a high of .32; the average construct convergence varied from –.14 to .19; and the correlation among specific construct pairs (not shown) ranged from a low of –.39 (*HVI* with *CYN*) to a high of .43 (*HVI* with Scale *6*).

The statistical theory of sampling distributions predicts that the average of the values observed across randomly selected samples should converge on the population parameter (Hays, 1981). In this instance, across the 300 samples, the average value for Variable A should have converged on the population parameter of .04, whereas the average value for Variable B should have been .03. This was the pattern obtained. In Figure 2, the average values were .03 and .02 for Variables A and B, respectively. In Figure 3, the values were precisely on target, being .04 and .03, respectively.

In general, these findings underscore how relatively small samples can produce results that differ dramatically from investigator to investigator. Due to natural fluctuations in sampling error, some samples will find that the Rorschach and MMPI first factors are positively correlated, while others will find they are negatively correlated or uncorrelated. Similarly, some samples (like the high *R* group in Meyer, 1993) will find Rorschach and MMPI construct pairs that just happen to be positively correlated, while other samples will find the same variables to be uncorrelated or negatively correlated (like the low *R* group in Meyer, 1993). When considering the latter points, it should be noted that the observed values for Meyer's (1993) initial *R* groups (i.e., the squares in Figure 1) fit within the parameters of the cloud of random error circles in Figure 2. In addition, if one took the observed values from the expanded *R* groups (i.e., the inverted triangles in Figure 1) and superimposed them over Figure 3, it would be seen that they also fit within the parameters of the random error cloud for samples of $n = 75$.

Figures 2 and 3 also demonstrate that the samples deliberately selected on the basis of response–character styles anchor the tails of each scattergram. Because these deliberately selected samples each contain a relatively large number of patients, Figure 3 provides the best indication of how these samples differ from randomly selected samples of patients. The data clearly indicate that selecting patients on the basis of these criteria is an optimal way to produce positive (or negative) convergent validity coefficients between Rorschach and MMPI constructs. The data also suggest relatively large samples are unlikely to provide substantial correlations between Rorschach and MMPI constructs unless: (a) patients are deliberately selected on the basis of response–character styles or (b) there are factors operating in the assessment environment that serve to align these stylistic qualities across methods.

Importantly, the extent to which Rorschach and MMPI first factors are correlated also helps determine the extent to which some meaningless variable pairs are correlated. Examination of the CUVP–*HighFF* rows in Table 5 reveals that even though the average coefficients were not statistically significant, they had a positive sign when patients had similar response–character styles and a negative sign when patients had opposing styles. This indicates that these meaningless variable pairs also tended to track first factor alignment. When there was a positive correlation between the first factors, there was a positive correlation for the average CUVP–*HighFF* coefficient; when there was a negative correlation between the first factors, there was a negative correlation for the average of these conceptually unrelated coefficients. As a result, when the average of the CUVP–*HighFF* was used as Variable B and the random draw analyses repeated, there continued to be a correlation between first factor alignment and the revised Variable B. This was true for samples of 30 patients ($r = .53$, $N = 300$, $p < .001$) and samples of 75 ($r = .44$, $N = 300$, $p < .001$). The key distinction between these analyses and those reported in Figures 2 and 3 was in the magnitude of the Variable B correlations; genuine construct pairs produced larger coefficients than meaningless construct pairs. A similar process was not evident for the CUVP–*NotFF* set. In fact, when the average of these meaningless variable pairs was used as Variable B and the random draw analyses repeated, there tended to be a negative correlation between first factor alignment and the revised Variable B (for 30 patients, $r = -.11$, $N = 300$, $p = .064$; for 75 patients, $r = -.13$, $N = 300$, $p = .021$). Overall then, these analyses indicate that when Rorschach and MMPI–2 first factors are aligned in a sample, there will tend to be a correlation between any variable pairs, so long as both members of the pair are highly correlated with their first factors. However, as indicated by Table 5, the magnitude of the convergent correlations is most pronounced with conceptually meaningful variable pairs.

Finally, it should also be noted that, although the data in Figures 2 and 3 demonstrate unequivocally that there is a strong relation between first factor correlations and construct correlations, the relation is far from perfect. Other factors contribute to the convergence and divergence of Rorschach and MMPI data. These other factors most likely include complexities in personality organization, the distinct types of characteristics that are effectively measured by each instrument (i.e., consciously recognized characteristics vs. implicit characteristics) and other sources of measurement error (e.g., unreliability, incomplete construct definition).

CONCLUSIONS

Several conclusions can be drawn from these analyses. First, *F* and *K* from the MMPI–2 and *R* and *Lambda* from the Rorschach are slightly less optimal ways to define response–character styles than the first factors from each method. Nonetheless, they were still relatively effective criteria for moderating Rorschach and MMPI–2 convergent validity—particularly for analyses of similar response–character styles. Furthermore, because the group of patients identified by these profile scores were largely distinct from the group of patients identified by the factor criteria (κs = .45 and .30), the convergent validity coefficients generated by the profile scores provide semi-independent support for the overall hypotheses regarding MMPI–2 and Rorschach interrelations (Meyer, 1996, 1997b). The phenomenon under consideration is strong enough so that substantially different operational definitions of response–character styles produced the same general results. Again, the findings were stronger and more differentiated for patients who displayed similar styles of interacting with the Rorschach and the MMPI than for patients who displayed opposing styles on each test.

Taken together, these findings draw attention to the contextual factors that are part and parcel of how testing data are gathered. To the extent that circumstances associated with the testing process lead to positive correlations between MMPI–2 and Rorschach first factors, positive correlations should be evident between certain MMPI–2 and Rorschach scales. To the extent that contextual factors generate negative correlations between response–character styles, certain construct scales should be negatively correlated. Thus, patients who volunteer for an ongoing research project may be as open or "dilated" on all testing methods as their character structures will allow. This may produce positive correlations between the MMPI and Rorschach stylistic factors and, in turn, between meaningfully related constructs. The strong convergent correlations obtained by Perry, Viglione, and Braff (1992) may reflect an example of this. Alternatively, if testing was conducted in a hospital where some patients were actively seeking discharge and others actively seeking admission, the contextual factors may result in more patients having motivated distortions in their test data (particularly Styles 1–M and 4–M) and these contingencies may generate a negative correlation between Rorschach and MMPI stylistic factors. If so, then one could expect negative correlations among some construct scales in this type of sample. Many other situational factors may influence response–character styles, including those found during forensic evaluations or the ways in which some healthy research volunteers may playfully produce dramatic and pathological-looking Rorschach responses (among others).

Intrapsychic factors within particular samples of patients may also serve to align first factor variance on the MMPI and Rorschach. For instance, patients with borderline personality disorder often produce dilated profiles on both the MMPI and the Rorschach (Exner, 1985; Gartner, Hurt, & Gartner, 1989; Zalewski & Archer, 1991), whereas other types of patients have defensive structures that produce opposing stylistic patterns on each method (Meyer, 1997b). To the extent that certain types of patients predominate in a sample, they may produce a particular degree of correlation among test interaction styles and also among trait constructs.

Systematically attending to such contextual factors will make assessment research more complicated—in part because it forces researchers to recognize that "validity" is not an inherent and static quality of individual test scales. Rather, the validity of individual test scores is also a function of how patients interact with the testing tasks. To provide accurate information, scores must be interpreted in light of response–character styles and the other idiographic information that accompanies a psychological assessment (i.e., reasons for the evaluation, psychological and symptomatic history, external motivations impinging on the assessment, observed behavior, scores on other test scales and other assessment methods, etc.). To the extent that research can begin to incorporate these additional complexities, the ensuing results will be more relevant to clinicians because these variables always affect the clinical interpretation of assessment data.

Clinicians have historically interpreted assessment data while considering such test-taking factors, even though appropriately sophisticated research has lagged behind. Simultaneously, the context-sensitive idiographic judgments made by clinicians are affected by many sources of bias and error (Arkes, 1981; Garb, 1994; Spengler, Strohmer, Dixon, & Shivy, 1995) that do not impinge on research results. It seems then the challenge for advancing personality assessment is to temper the contextual "error" of method variance that pervades research by designing studies that treat all scores in a manner that more directly parallels what is done in actual clinical practice, while simultaneously producing research that helps clinicians to temper the error that naturally affects applied clinical judgment.

On a separate point, it must also be recognized that the convergent validity coefficients reported in Tables 4 and 5 are not true indications of heteromethod validity. Rather, these coefficients indicate the extent of construct convergence that is evident when the Rorschach is made to function as if it were another self-report instrument (or, conversely, when the MMPI is made to function as if it were a Rorschach-like method). When two scales from the same method family are correlated, it always results in a confounding of genuine trait variance and systematic error variance. The analyses reported here do not remove the contributions of systematic error from the Rorschach and MMPI–2 scales. Rather, the analyses only aligned this error variance across methods—just as it is always aligned when scales from the same method family are correlated. As a result, to some unknown extent, the coefficients presented here overestimate the extent to which the MMPI–2 and Rorschach measure synonymous trait constructs. Analyses with the two sets of CUVPs indicated there is meaningful construct convergence that goes beyond methodological artifact, but alternative research strategies will be required to define the extent of construct overlap more precisely.

In summary, the analyses presented here lead to two primary conclusions. First, my statement that simple response frequency may moderate Rorschach–MMPI relations was incorrect (Meyer, 1993). Although it would have been wise to resist the urge to interpret those unexpected findings, efforts to make sense of the data eventually led to what I believe is a more refined understanding of some factors that generate empirical associations between Rorschach and MMPI–2 scores (Meyer, 1996, 1997b). A skeptic could reasonably argue that this process of theoretical evolution may reflect nothing more than yet another conceptual dead-end in the history of thought regarding Rorschach and MMPI convergence. That may be so. However, one can have some confidence in the more recent theorizing because it neatly explained why the initial findings related to *R* emerged in one study but did not replicate in an expanded sample of patients.

Second, it does seem warranted to conclude that response–character styles contribute to the convergent validity of MMPI–2 and Rorschach scales in this particular sample of adult patients. Even when substantially different groups of patients were selected—first by factor scores and then by profile scores—strong incremental construct convergence was evident, particularly for patients with similar test interaction styles (e.g., Tables 4 and 5). Furthermore, the randomly selected samples portrayed in Figures 2 and 3 provide good support for the hypothesis that, in general, the convergence of MMPI and Rorschach constructs is tied to broad stylistic qualities related to how patients interact with both assessment methods. These findings need to be replicated in other samples in order to understand the extent to which they generalize (cf. Archer, 1996). The research by Krishnamurthy et al. (1996) was a step in this direction. Unfortunately, the criteria they devised for defining response–character styles were quite different from the prior adult criteria. Thus, their findings should not be viewed as a failure to validate the prior research. Rather, confirmation, tempering, or negation of the findings will have to await further investigation. It would be more optimal to use the original factor criteria in such efforts, although the profile scores should also facilitate this goal.

ACKNOWLEDGMENTS

I thank Radhika Krishnamurthy and Ron Ganellen for their helpful comments regarding the issues presented here. I also thank Sharon Rae Jenkins, Bob Lovitt, and two reviewers for their helpful comments on earlier versions of this article.

REFERENCES

Adams, H. B., & Cooper, G. D. (1961). Rorschach response productivity and overt psychiatric symptomatology. *Journal of Clinical Psychology, 17,* 355–357.

Alterman, A. I., Snider, E. C., Cacciola, J. S., Brown, L. S., Jr., Zaballero, A., & Siddiqui, N. (1996). Evidence for response set effects in structured research interviews. *Journal of Nervous and Mental Disease, 184,* 403–410.

Archer, R. P. (1992). *MMPI–A: Assessing adolescent psychopathology.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Archer, R. P. (1996). MMPI–Rorschach interrelationships: Proposed criteria for evaluating explanatory models. *Journal of Personality Assessment, 67,* 504–515.

Archer, R. P., & Krishnamurthy, R. (1993a). Combining the Rorschach and the MMPI in the assessment of adolescents. *Journal of Personality Assessment, 60,* 132–140.

Archer, R. P., & Krishnamurthy, R. (1993b). A review of MMPI and Rorschach interrelationships in adult samples. *Journal of Personality Assessment, 61,* 277–293.

Archer, R. P., & Krishnamurthy, R. (1997). MMPI–A and Rorschach indices related to depression and conduct disorder: An evaluation of the incremental validity hypothesis. *Journal of Personality Assessment, 69,* 517–533.

Arkes, H. R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. *Journal of Consulting and Clinical Psychology, 49,* 323–330.

Belmont, I., & Birch, H. G. (1962). "Productivity" and mode of functioning in the Rorschach responses of brain-damaged patients. *Journal of Nervous and Mental Disease, 134,* 456–462.

Ben-Porath, Y. S., & Sherwood, N. E. (1993). *The MMPI–2 Content Component Scales: Development, psychometric characteristics, and clinical application.* Minneapolis: University of Minnesota Press.

Block, J. D. (1965). *The challenge of response sets: Unconfounding meaning, acquiescence, and social desirability in the MMPI.* New York: Appleton-Century-Crofts.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Manual for the restandardized Minnesota Multiphasic Personality Inventory: MMPI–2. An administrative and interpretive guide.* Minneapolis: University of Minnesota Press.

Butcher, J. N., & Williams, C. L. (1992). *Essentials of MMPI–2 and MMPI–A interpretation.* Minneapolis: University of Minnesota Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). San Diego, CA: Academic.

Cole, D. A., Howard, G. S., & Maxwell, S. E. (1981). Effects of mono- versus multiple-operationalization in construct validation efforts. *Journal of Consulting and Clinical Psychology, 49,* 395–405.

Cole, D. A., Truglio, R., & Peeke, L. (1997). Relation between symptoms of anxiety and depression in children: A multitrait–multimethod–multigroup assessment. *Journal of Consulting and Clinical Psychology, 65,* 110–119.

Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68,* 1152–1162.

Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. (1975). *An MMPI handbook: Vol. II. Research applications.* Minneapolis: University of Minnesota Press.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research.* New York: Dryden.

Edwards, A. L. (1964). Social desirability and performance on the MMPI. *Psychometrika, 29,* 295–308.

Edwards, L. K., & Edwards, A. L. (1991). A principal-components analysis of the Minnesota Multiphasic Personality Inventory factor scales. *Journal of Personality and Social Psychology, 60,* 766–772.

Epstein, S., Lundborg, E., & Kaplan, B. (1962). Allocation of energy and Rorschach responsivity. *Journal of Clinical Psychology, 18,* 236–238.

Exner, J. E., Jr. (1985). Some Rorschach data comparing schizophrenics with borderline and schizotypal personality disorders. *Journal of Personality Assessment, 50,* 455–471.

Exner, J. E., Jr. (1993). *The Rorschach: A comprehensive system: Vol. 1. Basic foundations* (3rd ed.). New York: Wiley.

Finch, A. J., Jr., Imm, P. S., & Belter, R. W. (1990). Brief Rorschach records with children and adolescents. *Journal of Personality Assessment, 55,* 640–646.

Finn, S. E. (1996). Assessment feedback integrating MMPI–2 and Rorschach findings. *Journal of Personality Assessment, 67,* 543–557.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.
Ganellen, R. J. (1994). Attempting to conceal psychological disturbance: MMPI defensive response sets and the Rorschach. *Journal of Personality Assessment, 63,* 423–437.
Ganellen, R. J. (1996). Exploring MMPI–Rorschach relationships. *Journal of Personality Assessment, 67,* 529–542.
Garb, H. N. (1994). Cognitive heuristics and biases in personality assessment. In L. Heath, R. S. Tindale, J. Edwards, E. Posavac, F. Bryant, E. Henderson, Y. Suarez-Balcazar, & J. Myers (Eds.), *Applications of heuristics and biases to social issues* (pp. 73–90). New York: Plenum.
Gartner, J., Hurt, S. W., & Gartner, A. (1989). Psychological test signs of borderline personality disorder: A review of the empirical literature. *Journal of Personality Assessment, 53,* 423–441.
Glutting, J. J., Oakland, T., & Konold, T. R. (1994). Criterion-related bias with the Guide to the Assessment of Test-Session Behavior for the WISC–III and WIAT: Possible race/ethnicity, gender, and SES effects. *Journal of School Psychology, 32,* 355–369.
Graham, J. R. (1993). *MMPI–2: Assessing personality and psychopathology* (2nd ed.). New York: Oxford.
Greene, R. L. (1991). *The MMPI–2/MMPI: An interpretive manual.* Boston: Allyn & Bacon.
Harkness, A., McNulty, J., & Ben-Porath, Y. (1995). The Personality Psychopathology Five (PSY–5): Constructs and MMPI–2 scales. *Psychological Assessment, 7,* 104–114.
Hays, W. L. (1981). *Statistics* (3rd ed.). New York: Holt, Rinehart & Winston.
Jackson, D. N., Fraboni, M., & Helmes, E. (1997). MMPI–2 content scales: How much content do they measure? *Assessment, 4,* 111–117.
Johnson, J. H., Butcher, J. N., Null, C., & Johnson, K. N. (1984). Replicated item level factor analysis of the full MMPI. *Journal of Personality and Social Psychology, 47,* 105–114.
Jourard, S. M. (1961). Self-disclosure and Rorschach productivity. *Perceptual and Motor Skills, 13,* 232.
Krishnamurthy, R., Archer, R. P., & House, J. J. (1996). The MMPI–A and Rorschach: A failure to establish convergent validity. *Assessment, 3,* 179–191.
Masling, J. (1960). The influence of situational variables and interpersonal variables in projective testing. *Psychological Bulletin, 57,* 65–85.
Messick, S., & Jackson, D. N. (1961). Acquiescence and the factorial interpretation of the MMPI. *Psychological Bulletin, 58,* 299–304.
Meyer, G. J. (1992a). Response frequency problems in the Rorschach: Clinical and research implications with suggestions for the future. *Journal of Personality Assessment, 58,* 231–244.
Meyer, G. J. (1992b). The Rorschach's factor structure: A contemporary investigation and historical review. *Journal of Personality Assessment, 59,* 117–136.
Meyer, G. J. (1993). The impact of response frequency on Rorschach constellation indices and on their validity with diagnostic and MMPI–2 criteria. *Journal of Personality Assessment, 60,* 153–180.
Meyer, G. J. (1994, May). *On the integration of assessment methods: The Rorschach and MMPI–2.* Paper presented at the 29th Annual Symposium on Recent Developments in the Use of the MMPI/MMPI–2/MMPI–A, Minneapolis, MN.
Meyer, G. J. (1996). The Rorschach and MMPI: Toward a more scientifically differentiated understanding of cross-method assessment. *Journal of Personality Assessment, 67,* 558–578.
Meyer, G. J. (1997a). Assessing reliability: Critical corrections for a critical examination of the Rorschach Comprehensive System. *Psychological Assessment, 9,* 480–489.
Meyer, G. J. (1997b). On the integration of personality assessment methods: The Rorschach and MMPI–2. *Journal of Personality Assessment, 68,* 297–330.
Meyer, G. J. (1998). Error in research and assessment data with an erratum for Meyer (1993). *Journal of Personality Assessment, 72,* 195–211.
Nichols, D., & Greene, R. (1997). Dimensions of deception in personality assessment: The example of the MMPI–2. *Journal of Personality Assessment, 68,* 251–266.

Oakland, T., & Glutting, J. J. (1990). Examiner observations of children's WISC–R test-related behaviors: Possible socioeconomic status, race, and gender effects. *Psychological Assessment, 2,* 86–90.
Perry, W., Viglione, D. J., Jr., & Braff, D. (1992). The Ego Impairment Index and schizophrenia: A validation study. *Journal of Personality Assessment, 59,* 165–175.
Pogge, D. L., Stokes, J. M., Frank, J., Wong, H., & Harvey, P. D. (1997). Association of MMPI validity scales and therapist ratings of psychopathology in adolescent psychiatric inpatients. *Assessment, 4,* 17–27.
Raghunathan, T. E., Rosenthal, R., & Rubin, D. B. (1996). Comparing correlated but nonoverlapping correlations. *Psychological Methods, 1,* 178–183.
Rappaport, D., Gill, M. M., & Schafer, R. (1968) *Diagnostic psychological testing* (Rev. ed.). New York: International Universities Press.
Robins, R. W., & John, O. P. (1997). Effects of visual perspective and narcissism on self-perception: Is seeing believing? *Psychological Science, 8,* 37–42.
Shatin, L. (1958). The constriction–dilation dimension in Rorschach and TAT. *Journal of Clinical Psychology, 14,* 150–154.
Shedler, J., Mayman, M., & Manis, M. (1993). The *illusion* of mental health. *American Psychologist, 48,* 1117–1131.
Spengler, P. M., Strohmer, D. C., Dixon, D. N., & Shivy, V. A. (1995). A scientist practitioner model of psychological assessment: Implications for training, practice, and research. *Counseling Psychologist, 23,* 506–534.
Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245–251.
Tsujimoto, R. N., Hamilton, M., & Berger, D. E. (1990). Averaging multiple judges to improve validity: Aid to planning cost-effective clinical research. *Psychological Assessment, 2,* 432–437.
Welsh, G. S. (1956). Factor dimensions *A* and *R*. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic readings on the MMPI in psychology and medicine* (pp. 264–281). Minneapolis: University of Minnesota Press.
Wohl, J. (1957). A note on the generality of constriction. *Journal of Projective Techniques, 21,* 410–413.
Zalewski, C., & Archer, R. P. (1991). Assessment of borderline personality disorder: A review of MMPI and Rorschach findings. *Journal of Nervous and Mental Disease, 179,* 338–345.

Gregory J. Meyer
Department of Psychology
University of Alaska Anchorage
3211 Providence Drive
Anchorage, AK 99508
E-mail: afgjm@uaa.alaska.edu

Received May 9, 1997
Revised September 28, 1998