

This is the author's final version of the contribution published as:

Viglione D; Giromini L; Gustafson ML; Meyer G. Developing Continuous Variable Composites for Rorschach Measures of Thought Problems, Vigilance, and Suicide Risk. *ASSESSMENT*. 21 (1) pp: 42-49.
DOI: 10.1177/1073191112446963

The publisher's version is available at:

<http://asm.sagepub.com/cgi/doi/10.1177/1073191112446963>

When citing, please refer to the published version.

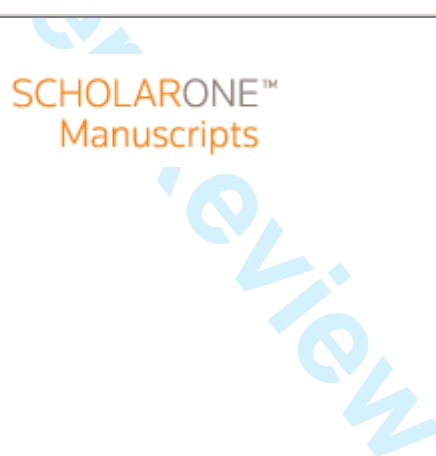
Link to this full text:

<http://hdl.handle.net/2318/1508195>



**Developing Continuous Variable Composites
for Rorschach Measures of Thought Problems,**

Journal:	<i>Assessment</i>
Manuscript ID:	ASMNT-12-0027
Manuscript Type:	Original Manuscript
Keywords:	Rorschach, Schizophrenia, Psychosis, Thought Disorder , Vigilance , Suicide, Composites



Abstract

Using a multiple regression approach with a large developmental sample (N = 460) of Rorschach protocols from psychiatric, forensic, and nonclinical control groups, we created continuous multi-variable Composite scores corresponding to the Comprehensive System (CS, Exner, 2003) Perceptual-Thinking Index (PTI), Hypervigilance Index (HVI), and Suicide Constellation (S-CON). Within a validation sample (N=230), these three new scores, called the Thought and Perception Composite (TP-Comp), Vigilance Composite (V-Comp), and Suicide risk-Composite (SC-Comp) were strongly associated with the three original CS Indices. Additional analyses suggest that the new Composite scores were more reliable than and at least as valid as the original Indices. Interpretive guidelines are offered.

Developing Continuous Variable Composites

for Rorschach Measures of Thought Problems, Vigilance, and Suicide Risk

To date, Rorschach systems have not taken full advantage of computers to perform complex calculations. In the Rorschach Comprehensive System (CS; Exner, 2003), for example, clinical Indices are derived from a series of dichotomous cut-off items that then lead to a step-based integer scale rather than a fully dimensional scale. Although these Indices can be calculated by hand, such integer-based scales, relative to dimensional, continuous scales, suffer from distributional and reliability problems because they assume only a limited number of values and fail to use all the available variance. For example, the Perceptual Thinking Index (PTI) is calculated with five individual algebraic statements or “steps” based on cut-off scores, each generating a value of 0 or 1, which are summed to create the total PTI score. One of the underlying steps is $M- > 1$ or $X-\% > .40$; if either is true, one point is added to the PTI value. Thus, the index can only assume integer values from zero to five. Decimal or intermediate values, such as 3.5 are not possible. One might describe this integer-based, discontinuous structure as having “cliffs.” These cliffs limit the reliability of the index in that small changes in a few responses could have a large impact on the index. As an extreme example, mis-coding a single human movement form quality minus response (M-) might add as much as three points to the PTI total by changing the outcomes for three steps that relate to form quality (i.e., using $XA\%$ and $WDA\%$ in Step 1, $X-\%$ in Step 2, and $X-\%$ or $M-$ in Step 3). Theoretically, one might expect limitations in the validity of the scale because cliffs cause it to be less highly correlated with the latent construct that it measures, thought and perceptual disturbance, which would be a demonstration of how reliability can limit validity. Support for this hypothesis also comes from a

1
2
3 recent review indicating that continuous measures of psychopathology are more valid and
4
5 reliable than discrete measures (Markon, Chmielewski & Miller, 2011). Thus, a fully
6
7 dimensional PTI could improve reliability and validity and thus also improve interpretive
8
9 accuracy in applied work.
10
11

12
13 The current study attempts to develop continuous variables corresponding to three of the
14
15 CS clinical Indices and to demonstrate basic reliability and validity for using these variables.
16
17 Based on available research findings (Meyer & Archer, 2001; Meyer, Hsiao, Viglione, Mihura,
18
19 & Abraham, 2012; Mihura, Meyer, Dumitrascu, & Bombel 2012; Viglione, 1999; Viglione &
20
21 Hilsenroth, 2001; Viglione & Meyer, 2008), we chose two Indices with the most empirical
22
23 support – the PTI and Suicide Constellation (S-CON) – and one Index that has some support and
24
25 a reasonable inference base related to coded behaviors that transparently generalize to extratest
26
27 behaviors – the Hypervigilance Index (HVI). By convention within the CS, the S-CON is
28
29 interpreted dichotomously as an indicator of suicide risk. The HVI is likewise dichotomously
30
31 interpreted as a focused, effortful, detail oriented, and vigilant cognitive style. However, research
32
33 (Lindgren & Carlsson, 2002; Meyer, 1997, 1999) has supported using the S-CON integer score
34
35 as a dimensional indicator of distress and the HVI integer score for wariness and paranoid states.
36
37 Thus, there is support for using them dimensionally as we do in this study rather than only as a
38
39 dichotomous indicator. Indeed, the latent constructs of thought and perceptual disturbance,
40
41 vigilance, and intensity of suicidal risk are dimensional in nature and thus are appropriately
42
43 measured on a continuous scale. To develop continuous variables corresponding to the PTI, HVI,
44
45 and S-CON, we substitute their dichotomous cut-scores and integer results with a regression-
46
47 based model that produces continuous scores. To distinguish from the CS Indices, we call these
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 new scores “Composites;” i.e., the Thought and Perception Composite (TP-Comp), Vigilance
4 Composite (V-Comp), and Suicide Concern Composite (SC-Comp).
5
6

7 8 Method 9

10 To produce the new Composites, the individual variables included in the steps of the
11 original CS Indices were used as predictors in multiple regression with the applicable CS clinical
12 Index as the criterion. For example, we used all of the variables found in the original five PTI
13 steps, X-%, M-, etc. as individual predictors with the PTI total score as the criterion variable.
14
15 The raw score (b weight) prediction equations were then used to calculate the Composites, which
16 were then tested for generalizability with an independent validation sample. The goal was to
17 maximize the association between the original CS Index and its corresponding dimensional
18 Composite score in the validation sample.
19
20
21
22
23
24
25
26
27
28

29 *Variable Selection* 30

31 Several modifications to the predictor variables from the original Indices were necessary.
32
33 Some of the original CS steps use a deviation score rather than a simple frequency. For instance,
34 S-CON Step 3 is positive if the Egocentricity Index is less than .31 or greater than .44. To create
35 a dimensional variant of this variable in the model, we used the absolute value of the deviation
36 from the mean of these two cut-points (i.e., $|X - .375|$). Similarly, for the S-CON criterion of $P <$
37 3 or $P > 8$, we used the absolute value of the deviation from the mean of these two cut-points
38 (i.e., $|X - 5.5|$). The Zd score has such limited empirical support (Meyer, Viglione, Mihura,
39 Erdberg, & Erard., 2011; Mihura et al., 2012) that its use is not advised. Given, that it is likely
40 not to be used consistently in the future, we replaced with a logical proxy based on the relative
41 degree of complexity evident in selecting and synthetically linking inkblot locations. More
42 specifically, we used the Location-Developmental Quality Complexity/R Index (Loc-DQ/R).
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Location and Developmental Quality are the central variables in producing Zd, so that a low Zd
4
5 score would correspond to a low Loc-DQ/R score. The fifth S-CON criterion contains a
6
7 deviation algorithm, $Zd < -3.5$ or $Zd > 3.5$; consequently, the absolute value of the difference
8
9 between Loc-DQ/R and its mean value in our developmental sample was used ($ILQ-compl/R -$
10
11 1.373).

12
13
14
15 In addition, some predictor variables were transformed to reduce their skew and kurtosis.
16
17 For example, distributions of FAB2, SumT, and MOR substantially diverged from a normal
18
19 distribution so transformations were applied to prevent the final formulae of the Composites
20
21 from being affected by nonnormality, excessive variability, and extreme outliers in any of the
22
23 individual variables. Said differently, we wanted to avoid allowing a single extreme value in any
24
25 of the individual variables to have an excessively large impact on the final score. In fact,
26
27 ultimately, we found that the congruence between the integer Indices and continuous Composites
28
29 at the interpretively important higher values was greatest when we took a very conservative
30
31 approach of transforming variables with skew values greater than one. In all but two cases the
32
33 square root transformation reduced the predictor variable skew sufficiently, but two variables
34
35 required an inverse reciprocal to the second power transformation $[-1/(x+1)^2]$; see Behrens, 1997).
36
37 Eventually, all variables included in the analysis had absolute skew values ≤ 1 and absolute
38
39 kurtosis values ≤ 1.5 , except for the transformed FAB2 variable (skew = 2.17, kurtosis = 2.83).
40
41
42
43
44

45 *Sample*

46
47
48 We compiled a large sample of computerized Rorschach protocols from our available
49
50 records. To fit with contemporary recommendations and data concerning administration to
51
52 optimize the range of responses obtained (Dean et al., 2007; Meyer et al., 2011), we used the
53
54 same procedures as those used to generate the R-PAS norms to statistically model the number
55
56
57
58
59
60

1
2
3 and distribution of responses in these protocols so they would closely approximate the
4
5 distribution of responses observed on each of the ten cards when using the R-Optimized (R-Opt)
6
7 administration method¹. This method encourages examinees to give 2 or 3 responses per card,
8
9 prompts for another response if only one is given to any card, and stops collecting responses
10
11 after the fourth to any card (see Meyer et al., 2011). Applying these R-Opt modeling procedures
12
13 to our pool of protocols left 1350 records. This sample included a substantial proportion of
14
15 records from non-clinical controls and non-patients or children. However, we wished to
16
17 maximize predictions for adults and clinical patients, given that the three Indices address clinical
18
19 issues and the S-CON is only calculated for adults. Consequently, all child records and most of
20
21 the control and nonpatient records were removed. The final sample included 690 records, 100 of
22
23 which were randomly selected from our pool of nonclinical controls or nonpatients. As expected,
24
25 the distribution for R in this sample ($M = 24.1$, $SD = 4.6$) closely approximated the normative
26
27 distribution for R ($M = 24.2$, $SD = 4.7$; Meyer et al., 2011).
28
29
30
31
32
33

34 Formulae for the Composites were developed within a randomly selected developmental
35
36 sample ($N = 460$) and tested with a validation sample ($N = 230$). These 690 adult Rorschach
37
38 records came from eight nonoverlapping subgroups, identified by patient status, evaluation
39
40 context, diagnosis, or presenting problem. 30.0% were from a mixed sample of inpatients and
41
42 outpatients evaluated for psychiatric reasons, 28.7% were from offenders, 14.5% from control or
43
44 non-patients, 9.1% from trauma patients, 4.1% from patients with schizophrenia or other
45
46
47
48
49

50
51 ¹ The R-Opt modeling procedure is described in detail in Meyer et al. (2011). Briefly, it allows
52
53 one to select suitable records collected using standard CS guidelines and randomly prune among
54
55 excess responses so as to exactly match the response level parameters observed in a sample of
56
57 123 target records collected by experienced examiners using R-Opt guidelines. Because R-Opt
58
59 greatly limits the number of cards with only one response, records with an insufficient number of
60
61 cards having more than one responses are dropped, such that the final pool of modeled records is
62
63 smaller than the original pool of available records.

1
2
3 psychotic disorders, 3.5% from patients with depression or substance abuse, 1.9% from forensic
4
5 psychiatric patients, and 8.3% from patients with other conditions. The same proportions of these
6
7 subgroups was retained in the developmental and validation samples.
8
9

10 Men comprise 56.6% of the sample and the mean age is 35.04 (SD = 11.54)². More than
11
12 50 examiners contributed protocols to the sample. Ethnicity is not well documented in some of
13
14 these archival, computerized files. However, in the largest subsample of mixed inpatients and
15
16 outpatients, almost one third of these individuals are African-American.
17
18

19 Results

20 We used the same strategy and method for calculating the three Composites. At first,
21
22 multiple regression equations were established within the developmental sample using the
23
24 original CS target Index as the dependent variable and fully dimensional versions of its
25
26 component scores as predictors. Non-significant predictors were removed and a new model was
27
28 tested. The prediction equation derived from the raw score b values produces a fully dimensional
29
30 total score to replace the original CS integer score. Each prediction equation was then used to
31
32 calculate this new dimensional scale in the derivation sample and subsequently in the
33
34 independent validation sample.
35
36
37
38
39

40 Descriptive data in Table 1 reveal that the distributions of the Composites closely
41
42 resemble the original Indices. Standard deviations are smaller for the all three Composites as a
43
44 result of regression to the mean. However, this reduction in variability also may be related to less
45
46 error in the Composites from replacing the individual integer steps with dimensional scores.
47
48

49 Because of the smaller SDs interpretive ranges to designate high, clinically relevant values will
50
51 be slightly lower for the Composite variables in contrast to the original CS Indices.
52
53

54
55 ² Six records were missing gender information and 34 were missing ages, but are known to be
56
57 adults.
58
59
60

TP-Comp

Table 2 provides information on the original Indices and their regression-based dimensional alternatives. When predicting the PTI, the transformed variable for Level 2 Special Scores [Lvl2; $(-1/(Lvl2+1)^2)$] was not a significant contributor and it was removed from the analysis. All other variables were significant. Within the developmental sample, the overall model was significant ($F(6,453) = 250.808, p < .001$), with a multiple $R = .88$. In the independent validation sample, the correlation between the TP-Comp and the observed PTI was .87. Distribution data in Table 1 reveal that the original PTI and new TP-Comp are highly similar. A supplemental table available from the first author provides the M, SD, and confidence interval for the Composite scores at each value of the original CS Index for both the developmental and the validation samples. These results illustrate the regression to the mean noted before. For instance, PTI values of 2, 3, 4, and 5 equate to lower TP-Comp values, with corresponding means of about 1.8, 2.5, 3.5, and 4.5.

V-Comp

All V-Comp predictor variables were significant and the model was significant ($F(8,451) = 171.372, p < .001$) with a multiple $R = .87$ in the developmental sample. The correlation between the V-Comp and the observed HVI within the validation sample was .86. As with TP-Comp, a close correspondence between the Index and Composite at all points of the original scale was achieved. Again, at the more extreme high HVI values, the corresponding V-Comp mean values are less, again demonstrating regression to the mean. HVI values of 5, 6, 7, and 8 have corresponding V-Comp means of approximately 4.5, 5.5, 6.0, and 7.0.

SC-Comp

1
2
3 R was not significant and was excluded from the regression analysis. Within the
4
5 developmental sample, the model was significant ($F(11,448) = 94.054, p < .001$) with a multiple
6
7 R of .84. The correlation with the observed S-CON in the validation sample was .79. Descriptive
8
9 data for the validation sample in Table 1 reveal nearly identical means. Because of regression to
10
11 the mean, S-CON values of 6, 7, 8, and 9 have corresponding SC-Comp means of roughly 5.7,
12
13 6.5, 7.0, and 8.0.
14
15

16 17 *Interrater Reliability*

18
19 Initial inter-rater reliability data for the Composites was recently investigated in a sample
20
21 of 50 Rorschach records collected with R-Opt administration (Viglione, Blume-Marcovici,
22
23 Miller, Giromini & Meyer, 2012). This sample was diverse in terms of age, sex, ethnicity,
24
25 educational background and patient status, and 16 examiners were involved. The exact
26
27 agreement intraclass correlations for a single rater (ICCs) for the Composites and Indices are
28
29 shown in Table 3 (upper part). To determine whether the interrater reliability for the new
30
31 Composites exceeded that of the corresponding original Indices we used procedures to compare
32
33 coefficients derived from the same sample that do not share overlapping variables (Raghunathan,
34
35 Rosenthal, & Rubin, 1996). The statistical test compares coefficients using the Pearson-Filon
36
37 statistic after they are transformed to Fisher's Z_r (ZPF). To obtain an effect size associated with
38
39 these comparisons, the difference in the Z_r coefficients was converted back from the Z_r metric to
40
41 the r metric. The resulting ZPF and r difference (r Diff) values shown in Table 3 indicate that the
42
43 reliability of the TP-Comp and V-Comp is superior than the reliability of the original Indices
44
45 (Table 3, upper part).
46
47
48
49
50
51

52
53 To test further the expected reliability gains with the Composites over the Indices, we
54
55 also re-analyzed two previously published data sets. In 2003, Viglione and Taylor reported
56
57
58
59
60

1
2
3 strong inter-rater reliabilities for CS variables in a mixed clinical and non-patient sample of 84
4 individuals. According to our new analyses, the interrater reliability for each of the Composite
5 scores falls in the excellent range and is significantly higher than the Index reliability (Table 3,
6 middle part). In 2002, Meyer et al. described interrater reliability for CS variables in several
7 samples collected by students (N = 66) and clinicians (N = 153). Given that the entire Meyer et
8 al.' sample size was large enough to select a subset of records with the R-Opt modeling
9 procedure, we computed the analyses both within the original and within the R-Opt modeled
10 samples (Table 3, lower part). Compared to the CS Indices, the Composites demonstrated higher
11 ICCs in all 18 comparisons, with significant differences ($p < .05$) in 13. Taken together, these
12 results indicate interrater reliability will be higher using the dimensional Composites rather than
13 the original CS Indices.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28

29 *Validation*

30
31 To investigate validity, we calculated the Indices and new Composites in previously
32 published studies. Dean et al. (2007) investigated the PTI among 61 residential care respondents,
33 31 of whom had psychotic disorders. These individuals were randomly assigned to either a CS
34 Rorschach administration or an alternative administration designed to control the number of
35 responses, similar to the R-Opt procedure. In the original study the PTI correlated at .381 ($p <$
36 .01) with a synthetic measure of thought disorder based on semi-structured interview, speech
37 sample, and self-report. The correlation between the thought disorder measure and the new TP-
38 Comp was .434 ($p < .01$). The difference between the two correlations was tested using
39 procedures for comparing correlations obtained from the same sample that share a common
40 variable (Meng, Rosenthal, & Rubin, 1992). This comparison was not statistically significant, Z
41 = 0.866, $p = 0.386$, but in the expected direction. In addition, when only the R-Opt like
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 alternative administration records ($n = 31$) were selected and the validity correlations re-
4
5 calculated, the TP-Comp seemed to be more valid than the PTI. Indeed, the PTI correlated .435
6
7 ($p < .05$), while the TP-Comp correlated .615 ($p < .001$); the difference between the two
8
9 correlations is statistically significant, $Z = 2.360$, $p = .018$, and of medium size, $r_{difference} =$
10
11 .246. In contrast, when only the standard CS administered records ($n = 30$) were selected, neither
12
13 the PTI ($r = .33$, $p > .05$) nor the TP-Comp ($r = .28$, $p > .05$) significantly correlated with the
14
15 measure of thought disorder. These results suggest that response distributions consistent with R-
16
17 Opt administration increases validity over standard CS administration, and within that context
18
19 TP-Comp has better validity than the PTI.
20
21
22
23

24
25 Another set of analyses for the PTI and TP-Comp was derived from a sample of 432
26
27 consecutive evaluations in a hospital-based assessment service in Chicago (Meyer, 2002). Data
28
29 has previously been reported on two proxy measures of psychotic functioning using diagnostic
30
31 information obtained from billing records before the psychological testing was initiated: (a)
32
33 diagnosis of a psychotic disorder and (b) severity of disturbance based on 5-point continuum
34
35 derived from the most severe diagnosis assigned (see Dawes, 1999; Meyer, 2002; Meyer &
36
37 Resnick, 1996; Viglione & Hilsenroth, 2001). Table 4 contains the correlations with these two
38
39 proxy criterion measures and shows both the correlations within the whole sample and within the
40
41 subset of R-Opt modeled records. In one of the four validity coefficient comparisons, the TP-
42
43 Comp was superior to the PTI, though with a small effect size.
44
45
46
47

48
49 Meyer (1997, 1999; Meyer, Riethmiller, Brooks, Benoit & Handler, 2000) reported on 87
50
51 clients selected from this same sample on the basis of concordant response style on the
52
53 Rorschach and MMPI-2. When test-taking response styles on the two tests was similar, as
54
55 measured by markers of each test's first factor, the CS Indices were substantially correlated with
56
57
58
59
60

1
2
3 corresponding MMPI-2 Indices. When response styles differed between the two tests,
4
5 correlations were either not significant or negative. Using patients with concordant test-taking
6
7 styles, Table 5 provides the correlation of the CS Indices with their corresponding MMPI-2
8
9 criteria reported in Meyer (1999), as well as, correlations for the corresponding Composites from
10
11 the current study, along with significance tests contrasting the two correlations³. Overall, these
12
13 correlations suggest equal validity for the clinical Indices and Composites. As explained in the
14
15 original articles, readers should keep in mind that these are artificially inflated coefficients using
16
17 just a subset of the original data after aligning method variance.
18
19
20
21

22 Discussion

23
24 Our aim was to develop continuous versions of the PTI, HVI, and S-CON in order to gain
25
26 reliability and precision over the cut-off-based integers used by the CS Indices. We used
27
28 dimensional versions of all the individual variables contained in the Indices and multiple
29
30 regression to create the continuous Composite variables. Using a relatively large developmental
31
32 sample (N = 460), we created the TP-Comp, V-Comp, and SC-Comp from the raw regression
33
34 coefficients applied to the individual predictor variables. These three continuous Composites
35
36 were tested in a validation sample (N=230), where the correlation between the original and new
37
38 score and the overall distribution parameters revealed a high degree of similarity. According to
39
40 these results, TP-Comp, V-Comp and SC-Comp are sufficiently identical to the three original CS
41
42 Indices that they can be substituted for them. Overall, our results suggest that the three
43
44 Composites have superior interrater reliability and are at least as valid as the original Indices. At
45
46
47
48
49
50
51
52
53
54

55 ³ In the original study, the Schizophrenia Index, a precursor of the PTI was reported, but not the
56 PTI itself. Accordingly, the PTI correlations, in addition to the Composite correlations were
57 calculated anew for this paper
58
59
60

1
2
3 the same time, however, it should be pointed out that these new Composites are statistically
4 complex and cannot be calculated by hand⁴, whereas their CS predecessors were not so limited.
5
6

7
8 As would be expected with regression toward the mean, extreme low values in the
9 original CS Indices correspond to slightly higher new Composite values, and extreme high
10 values in the original CS Indices correspond to slightly lower new Composite values. In terms of
11 clinical interpretations, however, this should not pose a problem as the results presented here
12 show that one should merely adjust the interpretive ranges from the original Indices to fit the
13 new, slightly narrower Composite distributions. Drawing on data from both the development and
14 validation samples in order to stabilize results coming from the tail ends of the distributions
15 where observations are not numerous, we can offer interpretive guidelines for the new
16 Composites by linking them to their equivalent scores on the original Index. In all instances, we
17 would expect Composite scores to be lower than Index scores. For instance, PTI values of 3, 4,
18 and 5 roughly correspond to TP-Comp values of 2.5, 3.5, and 4. Thus, a TP-Comp score of 3.5
19 suggests considerable concern about disturbed thinking and distorted perceptions of reality.
20 Previously, a PTI score of 3 has been interpreted as evidence of unusual thinking and perception.
21 TP-Comp scores approaching values of 2.5 should be interpreted in a similar fashion.
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 As noted earlier, the HVI in the CS is traditionally interpreted as dichotomous variable,
41 though Meyer (1999) and Lindgren and Carlsson (2002) found support for its use as a
42 dimensional variable. Moreover, it measures a focused, effortful, detail oriented, and vigilant
43 cognitive style. It thus requires some cognitive sophistication and psychological differentiation.
44 HVI values of 5, 6, 7, and 8 have corresponding V-Comp means of about 4.5, 5.5, 6.0, and 7.0.
45 From Table 1 it can be seen that a V-Comp score of 4.5 is almost one standard deviation above
46
47
48
49
50
51
52
53
54

55
56
57
58 ⁴ A file to calculate the Composites is freely available online (see Meyer et al., 2011).
59
60

1
2
3 the mean. Taking these data together, a V-Comp score exceeding 4.5 suggests a modest degree
4
5 of a focused, effortful, detail-oriented, and vigilant cognitive processing. A score of 5.5 suggests
6
7 a more extreme, inflexible, and vigilant cognitive style.
8
9

10 For the SC-Comp, a similar approach to anchoring interpretation is recommended.
11
12 From the literature an S-CON score of 7 suggests possible Suicide Concern that might or might
13
14 not be evident in self-report and a score of 8 indicates probable Suicide Concern (Fowler et al.,
15
16 2001; Viglione & Hilsenroth, 2001). S-CON values of 6, 7, 8, and 9 have corresponding SC-
17
18 Comp means of roughly 5.7, 6.5, 7.0, and 8.0. Thus, an SC-Comp score of 6.5 or above serves as
19
20 a suitable replacement for the S-CON score of 7, and would, therefore, suggest possible suicide
21
22 risk. We also recommend that a SC-Comp score of 7 be interpreted as suggesting a possible
23
24 Suicide Concern that is worthy of being evaluated, though with recognition of the risk for false
25
26 positive scores. A SC-Comp score of 8 should be interpreted as indicating a more worrisome risk
27
28 suicide.
29
30
31
32
33

34 The high similarity between the Indices and Composites indicates they are nearly
35
36 identical and should have the same network of relationships with important criteria.
37
38 Nevertheless, the validity findings reported here suggest the new TP-Comp is likely more valid
39
40 than the PTI, suggesting the continuous, dimensional distribution of the TP-Comp offers an
41
42 advantage in capturing more valid variance. The validity data for the V-Comp and SC-Comp
43
44 relied on more limited criteria and produced less conclusive results, suggesting that the validity
45
46 of these Composites approximates that of the original Indices. All three Composites clearly have
47
48 superior interrater reliability. Because of this better reliability, which comes from being
49
50 continuous rather than integer based, these Composites will likely be less affected by scoring
51
52
53
54
55
56
57
58
59
60

1
2
3 errors when used in the field, thus improving assessment accuracy in clinical and forensic
4
5 evaluations.
6

7
8 A limitation of this research is that the great majority of the records used across the
9
10 analyses were administered by CS methods and statistically modeled to approximate R-Opt
11
12 administration. However, our reliability sample obtained using R-Opt procedures performed just
13
14 like the modeled samples, suggesting that the findings should generalize across modes of
15
16 administration. This point is reinforced by seeing how similar the validity results are in Table 4
17
18 for the original Chicago protocols and for the protocols that remained after R-Opt modeling.
19
20 Finally, the strong validity support for the early R-Opt administration method used in the Dean et
21
22 al. (2007) data suggests that the TP-Comp findings should generalize to the R-Opt administration
23
24 procedure in R-PAS. Another limitation of this study concerns the type of psychopathology and
25
26 the patient status represented in these records, such that some populations may be over-
27
28 represented (e.g., forensic) and others under-represented (e.g., trauma). Accordingly, future
29
30 studies should test the validity of these scales with the R-Opt administration method, specific
31
32 clinical populations, and appropriate criteria. Also, in future research it would be worthwhile to
33
34 explore how simple z-score sums of variables might perform instead of using regression
35
36 weighted variables when creating new versions of these Indices.
37
38
39
40
41
42

43
44 Another avenue is opened by placing the V-Comp and SC-Comp on continuous scales so
45
46 that research can address the validity of these two Composites with dimensional criteria. For the
47
48 V-Comp, one can identify potential correlates associated with vigilance, paranoid processing,
49
50 and other cognitive and interpersonal phenomena. For the SC-Comp one could address
51
52 psychological issues such as suicide intentionality or despair. Also, determining whether there
53
54 are correlates to low scores on all three scales would be useful.
55
56
57
58
59
60

References

- 1
2
3
4
5
6 Behrens, J. T. (1997). Principles and procedures of exploratory data analysis. *Psychological*
7
8 *Methods*, 2, 131-160.
9
10 Dawes, R. M. (1999). Two methods for studying incremental validity of a Rorschach variable.
11
12 *Psychological Assessment*, 11, 297-302.
13
14
15 Dean, K., Viglione D.J., Perry, W., & Meyer G.J. (2007). A method to optimize response range
16
17 while maintaining Rorschach Comprehensive System validity. *Journal of Personality*
18
19 *Assessment*, 89(2), 149-161.
20
21
22 Exner, J. E. (2003). *The Rorschach: A comprehensive system*. (4th ed., Vol. 1: Basic foundations
23
24 and principles of interpretation). New York: Wiley.
25
26
27 Fowler, J. C., Piers, C., Hilsenroth, M. J., Holdwick, D. J., & Padawer, J. R. (2001). The
28
29 Rorschach Suicide Constellation: Assessing various degrees of lethality. *Journal of*
30
31 *Personality Assessment*, 76, 333-351.
32
33
34 Lindgren, T., & Carlsson, A. M. (2002). Relationship Between the Rorschach and the MMPI-2
35
36 in a Swedish Population: A Replication Study of the Effects of First-Factor Related Test-
37
38 Interaction Styles. *Journal of Personality Assessment*, 79, 357-370.
39
40
41 Markon, K. E., Chmielewski, M., & Miller, C. J. (2011). The reliability and validity of discrete
42
43 and continuous measures of psychopathology: A quantitative review. *Psychological*
44
45 *Bulletin*, 137(5), 856-879.
46
47
48 Meng, X. L., Rosenthal, R., & Rubin, D. B. (1992). Comparing correlated correlation
49
50 coefficients. *Psychological Bulletin*, 111, 172-175.
51
52
53 Meyer, G. J. (1997). On the integration of personality assessment methods: The Rorschach and
54
55 MMPI-2. *Journal of Personality Assessment*, 68, 297-330.
56
57
58
59
60

- 1
2
3 Meyer, G. J. (1999). The convergent validity of MMPI and Rorschach scales: An extension using
4
5 profile scores to define response-character styles on both methods and a re-examination
6
7 of simple Rorschach response frequency. *Journal of Personality Assessment*, 72, 1–35.
8
9
- 10 Meyer, G. J. (2002). Exploring possible ethnic differences and bias in the Rorschach
11
12 Comprehensive System. *Journal of Personality Assessment*, 78, 104-129.
13
14
- 15 Meyer, G. J., & Archer, R. P. (2001). The hard science of Rorschach research: What do we know
16
17 and where do we go? *Psychological Assessment*, 13, 486-502.
18
19
- 20 Meyer, G. J., & Resnick, G. D. (1996). *Assessing ego impairment: Do scoring procedures make*
21
22 *a difference?* Paper presented at the 15th International Research Conference, Boston.
23
24
- 25 Meyer, G. J., Hilsenroth, M. J., Baxter, D., Exner, J. E., Fowler, C. J., Piers, C. C., & Resnick, J.
26
27 (2002). An examination of interrater reliability for scoring the Rorschach Comprehensive
28
29 System in eight data sets. *Journal of Personality Assessment*, 78(2), 219–274.
30
31
- 32 Meyer, G. J., Hsiao, W.-C., Viglione, D. J., Mihura, J. L., & Abraham, L. H. (2012). A survey of
33
34 applied clinical validity for the Rorschach. Manuscript provisionally accepted for
35
36 publication at *Journal of Personality Assessment*.
37
38
- 39 Meyer, G. J., Riethmiller, R. J., Brooks, G. D., Benoit, W. A., & Handler, L. (2000). A
40
41 replication of Rorschach and MMPI-2 convergent validity. *Journal of Personality*
42
43 *Assessment*, 74, 175-215.
44
45
- 46 Meyer, G. J., Viglione, D. J., Mihura, J., Erard, R. E., & Erdberg, P. (2011). *Rorschach*
47
48 *Performance Assessment System: Administration, coding, interpretation, and technical*
49
50 *manual*. Toledo, OH: Rorschach Performance Assessment System.
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Mihura, J. L., Meyer, G. J., Dumitrascu, N., & Bombel, G. (2012). *A systematic review and*
4
5 *meta-analysis of the Rorschach Comprehensive System validity literature*. Manuscript re-
6
7 submitted for publication.
8
9
10 Raghunathan, Rosenthal, & Rubin (1996). Comparing correlated but nonoverlapping
11
12 correlations. *Psychological Methods, 1*, 178-183.
13
14
15 Viglione, D. J., & Hilsenroth, M. (2001). The Rorschach: facts, fiction, and future. *Psychological*
16
17 *Assessment, 13*, 452-471.
18
19
20 Viglione, D. J., & Meyer G. J. (2008). An overview of Rorschach psychometrics for forensic
21
22 practice. In C.B Gacono & F.B. Evans with N. Kaser-Boyd (Eds.) *Handbook of Forensic*
23
24 *Rorschach psychology*. (pp. 22-54). Matwah, NJ: Lawrence Erlbaum Associates.
25
26
27 Viglione, D. J., & Taylor, N. (2003). Empirical support for interrater reliability of the Rorschach
28
29 Comprehensive System coding. *Journal of Clinical Psychology, 59*(1), 111-121.
30
31
32 Viglione, D. J., Blume-Marcovici, A., Miller, H., Giromini, L. & Meyer G. J. (2012) An Initial
33
34 Inter-Rater Reliability Study for the Rorschach Performance Assessment System. (In
35
36 submission).
37
38
39 Viglione, D.J. (1999). A review of recent research addressing the utility of the Rorschach.
40
41 *Psychological Assessment, 11*, 251-265.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1

Descriptive Data for Responses, Original CS Indices, and New Dimensional Composites

Variable	Developmental (N=460)				Validation (N=230)			
	Mean R = 24.29 (SD = 4.68)				Mean R = 23.67 (SD = 4.51)			
	M	SD	Skew	Kurtosis	M	SD	Skew	Kurtosis
PTI	1.19	1.37	0.95	-0.12	1.26	1.30	0.83	-0.02
PT-Comp	1.19	1.21	0.66	0.52	1.28	1.20	0.68	1.17
HVI	3.33	1.61	0.20	-0.53	3.55	1.73	0.36	-0.62
V-Comp	3.33	1.39	0.17	0.02	3.50	1.50	0.20	0.18
S-CON	4.80	1.64	0.34	0.04	4.83	1.71	0.07	-0.43
SC-Comp	4.81	1.37	0.65	0.46	4.90	1.49	0.73	0.96

Table 2

Individual Variables Used in the Original CS Indices and in the Multiple Regression Equations to Create Dimensional Composite

Alternatives

Original Steps in the CS Indices	Alternative Variables in the Composite Equations	Final Regression Coefficients		
		Standardized β	raw b	p
PTI ^a				
<input type="checkbox"/> XA% < .70 & WDA% < .75	<i>Constant</i>		1.939	<.001
<input type="checkbox"/> X-% > .29	WDA%	-.217	-2.478	<.001
<input type="checkbox"/> Level 2 Special Scores > 2 & FAB2 > 0	X-%	.459	5.396	<.001
<input type="checkbox"/> R < 17 & WSum6 > .12 <i>or</i> R > 16 & WSum6 > 17	-1/(FAB2+1) ²	.119	.586	<.001
<input type="checkbox"/> M- > 1 <i>or</i> X-% > .40	R	-.055	-.016	=.022
	$\sqrt{\text{WSum6}}$.274	.186	<.001
	$\sqrt{\text{M-}}$.175	.335	<.001
HVI				
	<i>Constant</i>		.332	.272
<input type="checkbox"/> FT+TF+T = 0	$\sqrt{\text{SumT}}$	-.266	-.631	<.001
<input type="checkbox"/> Zf > 12	Zf	.199	.065	<.001
<input type="checkbox"/> Zd > 3.5	Loc_DQ-Complexity/R	.121	.699	<.001
<input type="checkbox"/> S > 3	S	.323	.224	<.001
<input type="checkbox"/> H+(H)+Hd+(Hd) > 6	$\sqrt{[\text{H}+(\text{H})+\text{Hd}+(\text{Hd})]}$.286	.652	<.001
<input type="checkbox"/> (H)+(A)+(Hd)+(Ad) > 3	$\sqrt{[(\text{H})+(\text{A})+(\text{Hd})+(\text{Ad})]}$.182	.388	<.001
<input type="checkbox"/> H+A : Hd+Ad < 4:1	$[\text{H}+(\text{H})+\text{A}+(\text{A})]/[\text{H}+(\text{H})+\text{A}+(\text{A})+\text{Hd}+(\text{Hd})+\text{Ad}+(\text{Ad})]$	-.208	-2.340	<.001
<input type="checkbox"/> Cg > 3	$\sqrt{\text{Cg}}$.134	.278	<.001
S-CON ^b				

Rorschach Composites 21

	Constant		2.020	<.001
□ FV+VF+V+FD>2	$\sqrt{(\text{SumV}+\text{FD})}$.157	.322	<.001
□ C. Sh Blends>0	$\sqrt{\text{C.Sh Blends}}$.351	.896	<.001
□ Ego<.31 or >.44	$\sqrt{ \text{Ego} - .375 }$.192	2.123	<.001
□ MOR>3	$\sqrt{\text{MOR}}$.129	.263	<.001
□ Zd>3.5 or <-3.5	$\sqrt{ \text{Loc_DQ-Complexity/R} - 1.373 }$.096	.882	<.001
□ Es>EA	es-EA	.196	.065	<.001
□ CF+C>FC	(CF+C)-FC	.244	.152	<.001
□ X+%<.70	X+%	-.078	-1.031	=.004
□ S>3	S	.251	.177	<.001
□ P<3 or >8	$\sqrt{ \text{P} - 5.5 }$.158	.589	<.001
□ PureH<2	H	-.156	-.128	<.001
□ R<17				

The final regression equations are derived by multiplying each variable by its raw b value, summing the results, and adding the constant. For instance, the TP-Comp score is obtained from this equation: TP-Comp = 1.939 - 2.478*(WDA%) + 5.396*(X-%) + .586*(-1/(FAB2+1)²) - .016*(R) + .186*($\sqrt{\text{WSum6}}$) +.335*($\sqrt{\text{MQ-}}$).

a. -1/(Lvl2+1)² was not significant and it was removed from the analysis; XA% was excluded because it is essentially the complement of X-%. In the developmental sample it was correlated with X-% at -.960 (p < .001);

b. R was not significant and it was removed from the analysis.

Table 3

Intraclass Correlation Coefficients for Interrater Reliability of the Dimensional Composites and Original Indices

Sample	Variable	ICC	ZPF	p	r Diff
Viglione et al. (2012) Sample; R-Opt Collected Data (N = 50) ^a .	PTI	.819	2.538	.011	.352
	TP-Comp	.909			
	S-CON	.809	.456	.648	.074
	SC-Comp	.833			
Viglione & Taylor (2003) Sample; Non Modeled Data (N = 84) ^b .	HVI	.918	3.289	.001	.462
	V-Comp	.969			
	PTI	.768	3.768	<.001	.381
	TP-Comp	.889			
Meyer et al. (2002) Sample; Non Modeled Data (N = 219) ^c .	S-CON	.876	2.418	.016	.291
	SC-Comp	.930			
	HVI	.947	4.586	<.001	.479
	V-Comp	.981			
Meyer et al. (2002) Sample; R-Opt Modeled Data (N = 93).	PTI	.917	1.105	.269	.074
	TP-Comp	.928			
	S-CON	.844	5.829	<.001	.381
	SC-Comp	.927			
Meyer et al. (2002) Sample; R-Opt Modeled Data (N = 93).	HVI	.910	6.051	<.001	.374
	V-Comp	.958			
	PTI	.885	2.740	.006	.275
	TP-Comp	.933			
Meyer et al. (2002) Sample; R-Opt Modeled Data (N = 93).	S-CON	.846	2.633	.008	.273
	SC-Comp	.909			
	HVI	.884	3.355	<.001	.371
V-Comp	.945				

Note: All individual ICCs are significant at $p < .001$.

a. ICCs for Composites were initially reported in Viglione et al. (2012)

b. ICCs for Indices were initially reported in Viglione & Taylor (2003)

c. ICCs for Indices were initially reported in Meyer et al. (2002)

Table 4

Composite Correlations with Psychosis and Diagnostic Severity in the Chicago Sample

	All Available Protocols—CS Administration				
	TP-Comp	PTI	Z	P	r Diff
Psychotic Diagnosis (N = 316)	.327	.275	2.026	0.043	0.057
Diagnostic Severity (N = 315)	.319	.283	1.400	0.161	0.040
	R-Opt Modeled Protocols				
	TP-Comp	PTI	Z	P	r Diff
Psychotic Diagnosis (N = 171)	.326	.312	0.404	0.686	0.016
Diagnostic Severity (N = 171)	.316	.309	0.202	0.840	0.008

Note: All correlations significant at $p < .0005$. Data derived from Meyer, 2002.

Table 5

Composite Correlations with Corresponding MMPI-2 Scales, N = 85

<i>Psychosis</i>					
<i>MMPI-2 Scales</i>	<i>PTI</i>	<i>TP-Comp</i>	<i>Z</i>	<i>p</i>	<i>r Diff</i>
Scale 8	.515***	.495***	-0.440	0.660	-0.027
BIZ	.444***	.452***	0.169	0.865	0.010
PSY-5-Psy	.453***	.457***	0.085	0.932	0.005
<i>Wariness</i>					
<i>MMPI-2 Scales</i>	<i>HVI</i>	<i>V-Comp</i>	<i>Z</i>	<i>p</i>	<i>r Diff</i>
Scale 6	.456***	.453***	-0.068	0.946	-0.004
CYN	.215*	.140	-1.537	0.124	-0.077
SOD	.386***	.375***	-0.241	0.810	-0.013
TRT2	.444***	.376***	-1.510	0.131	-0.082
<i>Affective Distress</i>					
<i>MMPI-2 Scales</i>	<i>SCON</i>	<i>SC-Comp</i>	<i>Z</i>	<i>p</i>	<i>r Diff</i>
Scale 2	.503***	.525***	0.360	0.718	0.030
Scale 7	.668***	.708***	0.800	0.424	0.076
DEP	.630***	.704***	1.435	0.151	0.133
ANX	.625***	.686***	1.163	0.245	0.107
PSY-5-Neg	.619***	.617***	-0.036	0.971	-0.003

a. Correlations were initially reported in Meyer (1999). However, two records with fewer than 14 responses were dropped so some of the coefficients differ slightly from those in the original publication.

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