Construct Validity of the MMPI–2 Restructured Clinical (RC) Scales: Reply to Rouse, Greene, Butcher, Nichols, and Williams

AUKE TELLEGEN,1 YOSEF BEN-PORATH,2 AND MARTIN SELLBOM2

1Department of Psychology, University of Minnesota
2Department of Psychology, Kent State University

Rouse, Greene, Butcher, Nichols, and Williams (2008) repeat two claims about the MMPI–2 Restructured (RC) scales. One asserts that the correlations of RC scales with parent Clinical scales are modest compared to the correlations with other existing MMPI–2 scales. In response, we reiterate that the RC scales were not meant to emulate the divergent and overlapping content of the Clinical scales. Instead, each represents a distinctive Clinical scale component. Although individually focused, the RC scales span collectively a wide range of content and used as multivariate predictors, account for most of the variance of each Clinical scale. Rouse et al. also claim that most RC scales are redundant with existing MMPI–2 scales, which they propose as substitutes (“proxies”). However, our analyses of Rouse et al.’s database and of our own data show that several of their proposed proxies are far less mutually distinguishable than are the RC scale counterparts. Furthermore, several Clinical scales are more successfully, and none are less successfully, accounted for by RC scales than by proxies. In response to Rouse et al.’s neglect of a body of empirical findings supporting the construct validity of the RC scales, we also review the relevant research literature.

In addition to stimulating corroborative empirical studies (see Ben-Porath & Tellegen, 2008, for a recent summary) and several positive reviews (e.g., Archer, 2006; Finn & Kamphuis, 2006; Simms, 2006; Weed, 2006), the MMPI–2 Restructured Clinical (RC) scales have also elicited negative reactions from persistent critics (e.g., Butcher, Hamilton, Rouse, & Cumella, 2006; Caldwell, 2006; Nichols, 2006). In this critique, Rouse, Greene, Butcher, Nichols, and Williams (2008) repeat two main objections, which readers may recognize from previous publications (e.g., Nichols, 2006). One is that the RC scales are too dissimilar from the Clinical scales to be acceptable as restructured versions. The second is that most RC scales are too similar to other existing MMPI–2 scales not to be redundant. Rouse et al.’s article focuses on marshalling evidence in support of the redundancy claim and selecting from among these existing scales the best RC scale substitutes (which we refer to as “proxies”). They make their selections on the basis of internal-structural evidence (scale intercorrelations and reliabilities).

Conspicuously absent from Rouse et al.’s (2008) critique is any consideration of the extensive body of empirical research findings documenting basic external RC scale correlates or any discussion of content characteristics. Yet, to adequately address the question of “what the RC scales measure,” in other words, to address the construct validity of the RC scales, an appraisal of not only internal-structural characteristics but also of external correlates and content features is essential.

To provide the necessary context for considering Rouse et al.’s (2008) questions and assertions about what the RC scales measure and for addressing the content of and lacunae in their arguments, we first recount briefly what the RC scales originally were and are still meant to measure. Two major and jointly compelling considerations motivated the effort to restructure the Clinical scales. One was the knowledge that the Clinical scales are valuable repositories of items empirically identified as capturing salient features of a wide range of major psychopathologies. The second is the longstanding recognition (e.g., Jackson, 1970; Norman, 1972) that as item composites (as aggregate measures), the Clinical scales warrant improvements.

The heterogeneity of the Clinical scales has been of particular concern. As was already explained some time ago (Nunnally, 1967) and has been reiterated by Tellegen et al. (2006) in response to Nichols’s (2006) previous critique, heterogeneous scales are not optimal for predictive purposes, including the prediction of complex syndromal criteria. The RC scales were created to capture important distinctive syndromal features embedded in the Clinical scales. From a contemporary psychometric perspective, this effort was overdue. Separate, nonoverlapping, and relatively homogeneous scales allow the predictive potential of the measured features to be maximized through familiar multivariate regression applications (for examples, see Tellegen et al., 2003, and Sellbom, Ben-Porath, & Graham, 2006).

In other words, the RC scales were not intended to emulate the heterogeneous makeup of the Clinical scales. Rather, each scale was designed to capture an important Clinical scale component. High correlations between the RC scales and the parent Clinical scales were not necessarily expected. However, although the RC scales are individually more focused than the Clinical scales, they cover collectively a wide range of content. Consequently, as we report shortly in more detail, the RC scales, when used
as multivariate predictors, account for most of the variance of each Clinical scale.

With respect to the question of redundancy, the reason for developing the RC scales was, as we just noted, to represent important Clinical scale components. Our objective was to arrive at a parsimonious set of substantively meaningful and theoretically promising measures tapping important aspects of major psychological disorders. As Tellegen et al. (2006) pointed out in response to earlier redundancy claims (Nichols, 2006), the methods used to construct the RC scales did not rule out the possibility that one or more of the scales would replicate existing measures. Nonredundancy was not a relevant consideration and was not an objective. Had an RC scale and a preexisting MMPI–2 scale been found to reflect a similar variance source, then differences in content quality or in psychometric efficiency could still have favored the RC scale. Even if the two scales had been found to be essentially interchangeable, this would not have diminished the value of the restructuring process as a method for identifying basic psychopathology constructs. So far, no such interchangeabilities have been encountered.

In sum, from a contemporary measurement viewpoint, an informative evaluation of whether and how the RC scales are related to the Clinical scales requires multivariate methods such as multiple regression. As for the issue of redundancy, correctly identified instances would be of interest even though they would have no bearing on the significance of the RC scales as measures of identified core psychopathology features. With these considerations in mind, we first address the two main questions Rouse et al. (2008) raised and believed to have answered: Are the RC scales too dissimilar from the MMPI–2 Clinical scales? Are the RC scales redundant with other preexisting MMPI–2 scales? To address these questions, we examine relevant internal-structural and content characteristics of the RC scales and of Rouse et al.’s proposed proxies. We also review the body of external correlates we referred to earlier, which were not addressed by Rouse et al. but are no less crucial to adequately appraising the construct validity of the RC scales and thus answering the question of what they measure.

For our answers, we draw on earlier reported findings as well as on additional analyses. To enhance comparability with Rouse et al.’s (2008) main findings (reported in their Tables 2–4), we used for these additional analyses a large mixed-gender sample, namely, the inpatient sample consisting of 501 men and 722 women described by Tellegen et al. (2003, chap. 4), which we refer to in the following as “our clinical sample.” For some of our analyses, we also examined two of Rouse et al.’s own samples.1 We selected two contrasting groups, one clinical (an inpatient sample), one non-clinical (a personnel screening sample): their “State Hospital” sample (234 men, 268 women) and their “Police” sample (6168 men, 926 women). Subdivided by gender, these two samples provided four subsamples. Individual scales will be identified by their abbreviated names (RC1, Clinical Scale 1, HEA, etc.). Full scale names and sources are shown in Table 1.

---

1We thank Steve Rouse for providing these data.

---

Although Rouse et al.’s (2008) empirical analyses focus on the issue of redundancy, in their introduction, they first call attention to the “marked contrast” between the high correlations of RC scales with preexisting MMPI–2 scales and the “more modest relationships between several RC scales and the parent [Clinical] scales from which a ‘distinctive core’ was distilled” (pp. 435–436). To illustrate this observation, Rouse et al. reproduced some of the correlations between RC scales and parent Clinical scales that Tellegen et al. (2003) reported for five subsamples. Perhaps not surprisingly, given the overall tenor of their critique, Rouse et al. identified the lowest correlations in Tellegen et al.’s (2003) tables. One of these, an unexpectedly low correlation of .26 between RC9 and Clinical Scale 9, happens to be in error;2 the correct value being .73. The other low correlation Rouse et al. identified, namely, between RC3 and Clinical Scale 3, has indeed consistently been the weakest one observed between an RC scale and its parent Clinical scale, ranging from .01 to −.20 in Tellegen et al.’s (2003) five clinical subsamples.

The apparent unrelatedness of RC3 to Clinical Scale 3 has made RC3 a favorite target of RC scale critics (Butcher et al., 2006; Caldwell, 2006; Nichols, 2006). Adopting the same univariate perspective that has characterized previous critiques, Rouse et al. (2008) linked Clinical Scale 3 exclusively to RC3 and arrived predictably at the same conclusion, namely, that no core characteristic of Clinical Scale 3 has been preserved.

We agree that RC3 is an especially instructive product of the restructuring effort but for a different reason. As Tellegen et al. (2006) noted before, important Scale 3 components are also represented in RC scales other than RC3. In Tellegen et al.’s (2003) five subsamples, the median correlation of Scale 3 with RC1 is .66 (range = .60–.76). This high correlation is attributable to the fact that somatic concerns are not only predominant in Clinical Scale 1 (and are therefore assessed by RC1) but are important in Clinical Scale 3 as well, if less overriding so because of the marked heterogeneity of Clinical Scale 3.

In our clinical sample, the correlation of RC1 with Clinical Scale 3 was .63, higher than any other RC Scale, particularly RC3, which correlated −.14 with Clinical Scale 3. However, to understand the predictive significance of both RC1 and RC3 from a multivariate perspective, in our clinical sample, we regressed Clinical Scale 3 on its best predictors among the nine RC scales as determined by a forward entry procedure. RC1 was of course selected first but was followed by RC3, which resulted in a multiple correlation of .76, with a beta weight of .81 for RC1 and of −.45 for RC3. In other words, combined with a second predictor variable, although not in isolation, RC3 was found to make a strong contribution to the prediction of Clinical Scale 3. Adding a third predictor raised the multiple correlation to .80. Essentially the same results (not reported here) were obtained with other clinical samples.

---

2The incorrect value of .26 reported for one of the correlations between RC9 and Clinical Scale 9 appeared in the initial printing of the Tellegen et al.’s (2003) monograph. Soon afterward, an erratum was mailed to all recipients of this first printing. Rouse et al. (2008) also incorrectly reported that the correlations between RC3 and Clinical Scale 3 range from −.13 to −.20. The actual correlations, accurately reported from the beginning, range from .01 to −.20.
In our clinical sample, we also used multiple regression to predict each of the other 7 Clinical scales from the RC scale set. In all solutions, each RC scale was again selected as either the first or second predictor of its parent Clinical scale. The zero-order correlations of each RC scale with its parent scale, followed by the multiple correlation based on three predictors, are as follows: Clinical Scale 1 (.95, .96); Clinical Scale 2 (.83, .89), Clinical Scale 3 (.83, .89), Clinical Scale 4 (.83), Clinical Scale 5 (.83), Clinical Scale 6 (.83), Clinical Scale 7 (.83, .89), Clinical Scale 8 (.83), Clinical Scale 9 (.83, .94), and Clinical Scale 10 (.94, .96).

The multiple correlations are all substantial but vary in magnitude. To appreciate these variations, it is important to take into account that Clinical Scales 2, 3, 4, 6, and 9, but not Clinical Scales 1, 7, and 8, include sizeable percentages (ranging from 32%–47%) of items once classified as “subtle” (Wiener & Harmon, 1946). According to two of the authors of Rouse et al.’s (2008) article, it is likely that these items ended up in the original Clinical scales by chance (Butcher & Williams, 2000), an opinion Butler (2005) has recently reaffirmed. The inclusion of seemingly subtle but actually invalid items would naturally be expected to lower the multiple correlations. Those we just reported for the five Clinical scales containing subtle items range from .80 to .89 (median = .83); those for the three without subtle items range from .94 to .96 (median = .96).

These results justify the conclusion that the RC scales account for a very large portion of the meaningful variance of each Clinical scale. The increments achieved when predictor scales are added demonstrate the extent to which Rouse et al.’s (2008) restrictive univariate approach, relating each Clinical scale exclusively to its one designated “offspring” scale, obscures meaningful connections linking Clinical scales to more than one Restructured scale.

### The RC Scale Set Is Not Redundant With and Cannot Be Replaced by Rouse et al.’s (2008) Set of Proposed Proxies

In the preceding section, we commented on Rouse et al.’s (2008) treatment of each RC scale as sole predictor of its parent Clinical scale. In response, in our clinical sample, we used more task-appropriate multivariate analyses allowing each Clinical scale to be predicted by a combination of RC scales. The results were informative, indicating that the RC scales appear close to achieving sufficiency as Clinical scale predictors.

However, in the main investigative portion of their article, Rouse et al. (2008) focused on the necessity of each RC scale as a predictor of its parent Clinical scale. In response, in our clinical sample, we used more task-appropriate multivariate analyses allowing each Clinical scale to be predicted by a combination of RC scales. The results were informative, indicating that the RC scales appear close to achieving sufficiency as Clinical scale predictors.

#### Table 1

<table>
<thead>
<tr>
<th>Scale Name</th>
<th>Abbreviation</th>
<th>Source</th>
<th>No. of RC Scale (Proxy Scale) Items</th>
<th>Item Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demoralization</td>
<td>RCd</td>
<td>Tellegen et al. (2003)</td>
<td>24 (A: 39)</td>
<td>14</td>
</tr>
<tr>
<td>Somatic Complaints</td>
<td>RC1</td>
<td>Tellegen et al. (2003)</td>
<td>27 (HEA: 36)</td>
<td>20</td>
</tr>
<tr>
<td>Low Positive Emotions</td>
<td>RC2</td>
<td>Tellegen et al. (2003)</td>
<td>17 (INTR: 34)</td>
<td>9</td>
</tr>
<tr>
<td>Cynicism</td>
<td>RC3</td>
<td>Tellegen et al. (2003)</td>
<td>15 (CYN: 23)</td>
<td>12</td>
</tr>
<tr>
<td>Antisocial Behavior</td>
<td>RC4</td>
<td>Tellegen et al. (2003)</td>
<td>22 (AAS: 13)</td>
<td>7</td>
</tr>
<tr>
<td>Ideas of Persecution</td>
<td>RC6</td>
<td>Tellegen et al. (2003)</td>
<td>17 (PSYC: 25)</td>
<td>10</td>
</tr>
<tr>
<td>Dysfunctional Negative Emotions</td>
<td>RC7</td>
<td>Tellegen et al. (2003)</td>
<td>24 (A: 39)</td>
<td>10</td>
</tr>
<tr>
<td>Aberrant Experiences</td>
<td>RC8</td>
<td>Tellegen et al. (2003)</td>
<td>18 (BIZ: 23)</td>
<td>11</td>
</tr>
<tr>
<td>Hypomanic Activation</td>
<td>RC9</td>
<td>Tellegen et al. (2003)</td>
<td>28 (Ho: 50)</td>
<td>7</td>
</tr>
<tr>
<td>Rouse et al. (2008) RC scale proxies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>A</td>
<td>Welsh (1956)</td>
<td>39 (RCd: 24, RC7: 24)</td>
<td>14, 10</td>
</tr>
<tr>
<td>Health Concerns</td>
<td>HEA</td>
<td>Butcher, Graham, Williams, &amp; Ben-Porath (1990)</td>
<td>36 (RC: 27)</td>
<td>20</td>
</tr>
<tr>
<td>Introversive/Low Positive Emotion</td>
<td>INTR</td>
<td>Harkness, McNulty, Ben-Porath, &amp; Graham (2002)</td>
<td>34 (RC2: 17)</td>
<td>9</td>
</tr>
<tr>
<td>Cynicism</td>
<td>CYN</td>
<td>Butcher et al. (1990)</td>
<td>23 (RC3: 15)</td>
<td>12</td>
</tr>
<tr>
<td>Addiction Admission Scale</td>
<td>AAS</td>
<td>Weed, Butcher, McKenna, &amp; Ben-Porath (1992)</td>
<td>13 (RC4: 22)</td>
<td>7</td>
</tr>
<tr>
<td>Psychoticism</td>
<td>PSYC</td>
<td>Harkness et al. (2002)</td>
<td>25 (RC6: 17)</td>
<td>10</td>
</tr>
<tr>
<td>Bizarre Mentation</td>
<td>BIZ</td>
<td>Butcher et al. (1990)</td>
<td>23 (RC8: 18)</td>
<td>11</td>
</tr>
<tr>
<td>Hostility</td>
<td>Ho</td>
<td>Cook &amp; Medley (1954)</td>
<td>50 (RC9: 28)</td>
<td>7</td>
</tr>
</tbody>
</table>

In the preceding section, we commented on Rouse et al.’s (2008) set of proposed proxies from a discriminant perspective. That is, we ask whether the proxy scales on the whole are more, equally, or less distinctive from
one another compared to the RC scales. We then assess the two sets of scales from a convergent perspective, namely, by comparing the multiple correlations we just reported, which were obtained when the RC scale set was used to predict each Clinical scale, with the multiple correlations obtained when Rouse et al.’s proxy set is used for that purpose.

Several of Rouse et al.’s (2008) RC Scale Proxies Lack Distinctiveness

To evaluate a set of predictor scales, it is informative to assess the degree of correlational independence between the scales. All other things equal, the more distinctive (the less highly intercorrelated) the predictor scales are, the greater their maximum combined predictive potential. Scale independence is of course not the only consideration. Certain clinically important characteristics (such as different manifestations of internalizing tendencies) tend to co-occur. To be optimal, measures assessing these tendencies may need to be similarly linked.

The sole criterion in Rouse et al.’s (2008) selection of proxies was a high correlation with the targeted RC scale. From all available MMPI-2 scales, Rouse et al. selected as proxy the one that in their 25-sample database was on average most strongly correlated with that RC scale. Rouse et al. did not take the correlations between potential proxies for different RC scales into account. This approach did not rule out the possibility that correlations between proxies for different RC scales are too high or too low to be satisfactory. For example, in our clinical sample, the correlation between the proxies for RC3 (CYN) and RC9 (Ho) is .92 compared to the correlation of .54 between RC3 and RC9, an increase of 55% in shared variance. Similarly, the proxies for RC6 (PSYC) and RC8 (BIZ) correlate .93 compared to the corresponding RC scale correlation of .69, a variance overlap increase of 39%.

To follow up on these observations, we used Rouse et al.’s (2008) four subsamples to examine correlational patterns in a more comprehensive manner. In each subsample, we compared each of the 36 squared correlations between the 9 RC scales with the squared correlation between its proxies. Figure 1 serves to facilitate these individual comparisons as well as the recognition of overall patterns. Each data point represents a pair of squared correlations: between two RC scales and between the two corresponding proxies. Any point located at or near the diagonal represents a pair of squared correlations that are (approximately) the same. Any point above the diagonal represents a squared correlation between two proxy scales that is higher than that between the two corresponding RC scales; any point below it depicts the opposite: a squared correlation between two RC scales higher than that between the two corresponding proxies. The farther removed a point is from the diagonal, the greater the discrepancy between the two squared correlations.

To illustrate, in Figure 1a, the data point labeled “68” identifies on the horizontal axis the square of the correlation of .61 (i.e., .37) between RC6 and RC8 and on the vertical axis the square of the correlation of .90 (i.e., .81) between the two proposed proxies for RC6 and RC8 (viz., PSYC and BIZ). Another data point, “d7,” identifies on the horizontal axis the square of the correlation of .79 (i.e., .62) between RCd and RC7 and on the vertical axis the (square of) correlation of 1.00 between Scale A and itself because Rouse et al. (2008) designated Scale A as a proxy for both RCd and RC7.

In each panel of Figure 1, only data points representing differences between squared correlations exceeding .15 in three of Rouse et al.’s (2008) four subsamples have been identified with labels. The labeled data points reveal a consistent pattern: All but one of the major differences between RC scale pairs and corresponding proxy pairs represent higher correlations between the proxies (are located above the diagonal). Several of these proxy pairs do not come close to matching the distinctiveness of the corresponding RC scales.

The one exception is the squared correlation between RCd and RC2, found to be higher (averaging .38 across Rouse et al.’s, 2008, four subsamples) than that between the two designated proxies, the A and INTR scales, respectively (averaging .20). In the next section, we explain why this finding is also a decidedly undesirable consequence of Rouse et al.’s (2008) proxy selection method. In sum, the patterns of discrepant correlations depicted in Figure 1 support the conclusion that from an internal-correlational perspective, Rouse et al.’s set of proxies does not replicate and cannot replace the RC scale set.

Rouse et al.’s (2008) Set of Proxy Scales Accounts for Less Clinical Scale Variance Than Does the RC Scale Set

We used Rouse et al.’s (2008) set of proxy scales to predict each Clinical scale in the same way we earlier reported using the RC scales for that purpose. With three proxies as predictors, we obtained, in our clinical sample, the following multiple correlations, each accompanied in parentheses by the multiple correlation achieved with the RC scale set: Clinical Scale 1: .95 (.96), Clinical Scale 2: .87 (.89), Clinical Scale 3: .80 (.80), Clinical Scale 4: .76 (.82), Clinical Scale 6: .76 (.87), Clinical Scale 7: .96 (.96), Clinical Scale 8: .94 (.94), and Clinical Scale 9: .74 (.83). These results show that the RC scales either match or exceed the predictive power of Rouse et al.’s proxies. For some Clinical scales (Scales 4, 6, and 9), the predictive advantage of the RC scales is appreciable. It is clear that for capturing Clinical scale variance, the performance of Rouse et al.’s set of proxies has not shown the RC scales set to be unnecessary and replaceable.

ADDITIONAL METHODOLOGICAL PROBLEMS WITH ROUSE ET AL.’S (2008) ANALYSES

Before examining Rouse et al.’s (2008) evaluations of each individual RC scale and its proxy, we identify a number of additional problems with several or all of these appraisals:

Problematic Interpretations of Reliability

Reliability and dimensionality. Rouse et al. (2008) interpreted a higher mean interitem correlation (IIC) coefficient to indicate that the scale in question is “more unidimensional.” It has been repeatedly pointed out that coefficient alpha (which in standardized form is a function of the mean IIC coefficient and number of items) is not an index of unidimensionality (e.g., Cortina, 1993; Crocker & Algina, 1986; Schmitt, 1996; Streiner, 2003). The same is true for the mean IIC. The mean IIC cannot be interpreted as an index of unidimensionality unless the variance of the individual IICs is relatively small. Conversely, variable IICs do not prove multidimensionality because even a strictly unidimensional scale may consist of items with variable factor loadings.
WHAT THE RC SCALES MEASURE

Reliability and validity. Rouse et al. (2008) asserted that “alpha is directly related to the power of a scale to correlate with other variables (Henson, 2001)” (p. 440). Although scale alphas and scale intercorrelations are not independent, the two statistics are not redundant. It is entirely possible that of two alternative predictor scales, the one with the lower alpha coefficient contains a larger valid variance component and is more highly correlated with relevant criterion variables.

For example, in Tellegen et al.’s (2003) three inpatient samples, the median alpha coefficients of RC8 and Clinical Scale...
8 are .85 and .94, respectively. Yet, in the same samples, the median correlations of the two scales with an intake-interview-based “hallucinations” rating scale are .37 and .15, respectively (and .15 and −.03, respectively, with the less central “delusions” rating). The higher reliability but lower external validity of Clinical Scale 8 is attributable to its greater length and greater saturation with Demoralization enhancing its alpha coefficient but its strong Demoralization component weakening it as a measure of thought disturbance. Specifically, in the three samples, the median correlation of Clinical Scale 8 with RCd is .82; the median correlation of RC8 with RCd is substantially lower, .45; whereas the median correlation of RCD with “hallucinations” is −.01. As this example demonstrates, reliability comparisons cannot substitute for the direct evidence on the comparative validities of the RC scales and the Clinical scales that Rouse et al. (2008) did not consider.

**Failure to Consider Scale-Overlap Artifacts**

Scale reliabilities, although not directly determining scale intercorrelations, set upper limits to these correlations. Scale intercorrelations that exceed reliability-imposed limits are suspect. Rouse et al.’s (2008) Tables 3 and 4 reveal that for seven of the nine scale pairs made up of an RC scale and its proxy, the average within-pair correlation exceeds the theoretical maximum as estimated from the average alpha coefficients calculated for the RC Scale and proxy in question. For four of the pairs (RC1/HEA, RC3/CYN, RC4/AAS, and RC8/BIZ), the reported shared variance exceeds the reliability-based maximum by 14% to 21%. For example, Rouse et al.’s Table 3 shows that the average correlation between RC1 and HEA equals .90. However, according to Rouse et al.’s Table 4, the average reliability of both scales equals only .82, implying a surplus of 14% shared variance (.90² − .82² = .14), an anomalous result.

Item overlap is undoubtedly one source of correlational inflation. For example, RC1 shares 20 of its 27 items with its 32-item proxy, the HEA Content scale. However, Rouse et al. (2008) seemed to reject overlap corrections on the mistaken assumption that such corrections are made by removing the overlapping items, and they criticize Tellegen et al. (2006) for having done so. Historically, this is what Welsh (1956) in effect did when he derived his A and R scales by factor analyzing “pure” (nonoverlapping) Clinical scales. However, this was done more than half a century ago. The current generally accepted method (the one actually used by Tellegen et al., 2006) is designed to remove only contributions of the spuriously perfect correlations between the random measurement error components of the overlapping portions of the two to-be-correlated scales (e.g., Hsu, 1992) and has been in place for some time (e.g., Tellegen & Briggs, 1967).

Using our clinical sample, we applied this correction to the case of RC1 and HEA, with the following results: (a) reliabilities of .870 and .882 and thus an estimated maximum correlation of \( r_{\text{max}} = \sqrt{(.870)(.882)} = .876 \); (b) an uncorrected correlation of .944 exceeding this value (and essentially the same as those Rouse et al., 2008, reported for their inpatient samples); and (c) a corrected correlation of .944 − .075 = .869, very close to but not exceeding the estimated maximum and thus itself a more plausible result than .944.

We believe that appropriately overlap-corrected estimates of the correlations between the RC scales and their designated proxies will allow more realistic appraisals of scale redundancy claims such as those Rouse et al. (2008) advanced.

**Repeating a Previously Refuted Criticism of the RC Scales**

Rouse et al. (2008) claimed that “A particularly thorny issue for use of the RC scales in clinical settings was the observation that the RC scales did not have elevations in the clinical range.” However, Tellegen et al. (2006, pp. 151–153) empirically demonstrated that the real problem resided in the use of Caldwell’s (1997) highly unrepresentative and heterogeneous collection of cases for examining RC scale properties in a supposedly “clinical” setting. Tellegen et al. (2006) showed that in representative clinical samples, RC scale and Clinical scale elevations are consistently comparable. Handel and Archer (2008) recently reported similar findings with a different clinical sample.

**WHAT DO THE RC SCALES MEASURE?**

As we noted earlier, a serious response to the question “What do the RC scales measure?” would have to deal with construct validity. It would address the substantive, internal-structural, and external-correlate characteristics of each scale (Loevinger, 1957). Pragmatic concerns, such as scale length, would not be neglected either.

So far, we have discussed general limitations in Rouse et al.’s (2008) internal-structural analyses. We have also noted their failure to consider external correlates. In their own words: “The data sets for this study did not allow for comparison of the RC scales and extant scales in terms of external correlates, and that type of data will be valuable for the ongoing evaluation of the RC scales” (p. 441). However, there is a substantial body of published empirical research, including a study published by Sellbom, Ben-Porath, et al. (2006) in this journal, directly comparing external correlates of the RC scales with a set of proxies very similar to the one they propose. Sellbom, Ben-Porath, et al. (2006) reported evidence of substantially stronger convergent validity for the RC scales.

Given the demonstrated lack of psychometric equivalence between the RC scales and Rouse et al.’s (2008) proxies, consequential differences between the external correlates of these two sets of scales are also possible. Therefore, we consider in the following not only content and internal-correlational evidence neglected by Rouse et al., but as an additional corrective, we also summarize, compare, and contrast the external correlate findings for each RC scale and its proxy.

**RCd.** Rouse et al. (2008) considered both RCd and its proposed proxy, Welsh’s A scale, to be measures of the “general factor” we have called Demoralization. Tellegen et al. (2003, 2006) have found Demoralization to be the MMPI–2 equivalent of the Pleasant–Unpleasant (or Happy–Unhappy) mood dimension in Watson and Tellegen’s (1985) widely studied model of affect. Consistent with this model, RCd is strongly associated with both (low) positive and negative temperament as measured in three different personality inventories: Tellegen’s (1982; Tellegen & Waller, 2008) Multidimensional Personality Questionnaire (Sellbom & Ben-Porath, 2005), Costa and McCrae’s (1992) Revised NEO Personality Inventory (NEO–PI–R), a measure of the Five-factor model of personality (Sellbom,
In keeping with the broadly unhappy content endorsed by individuals scoring high on Demoralization, external correlates of RCd include generalized emotional distress and unhappiness, depressed mood, and anxiety (Arbisi, Sellbom, & Ben-Porath, 2008; Forbey & Ben-Porath, 2007b, 2008; Handel & Archer, 2008; Osberg, Haseley, & Kamas, 2008; Sellbom, Ben-Porath, & Bagby, 2008a; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, & Schen, 2006; Tellegen et al., 2003, 2006). Congruent with these findings, RCd is also associated with symptoms of nonspecific distress such as decreased sleep, decreased appetite, guilt, sense of worthlessness, decreased energy, and poor concentration (Arbisi et al., 2008; Handel & Archer, 2008). Additional specific and clinically important symptoms and experiences associated with RCd include insecurity, interpersonal sensitivity, low self-esteem (Osberg et al., 2008; Sellbom, Ben-Porath, et al., 2006; Tellegen et al., 2003), life dissatisfaction, hopelessness and resulting loss of motivation, and suicidality (Arbisi et al., 2008; Forbey & Ben-Porath, 2007a; Handel & Archer, 2008; Tellegen et al., 2003, 2006; Wygant et al., 2007).

Corroborating its discriminant validity, RCd is more strongly correlated with the common component of distress disorders (depression, generalized anxiety disorder, and posttraumatic stress disorder) than with that of fear-based disorders (agoraphobia, social phobia, and specific phobias), which are more highly correlated with RC7 (Sellbom et al., 2008a; Tellegen et al., 2006). Furthermore, RCd is hardly or only weakly correlated with behavioral and thought dysfunction (Arbisi et al., 2008; Forbey & Ben-Porath, 2007a; Handel & Archer, 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Tellegen et al., 2003, 2006).

Findings such as those reported by Graham, Ben-Porath, and McNulty (1999) have shown a somewhat similar pattern of convergent and discriminant validity associations for Scale A. However, Tellegen et al. (2003, 2006) have demonstrated that RCd items are particularly strong first-factor markers. Furthermore, the empirical literature linking Scale A to a contemporary theoretical framework is limited. Although the two scales are similar in content, RCd, with 24 items, is appreciably shorter than the 39-item A scale. If brevity is a relevant practical concern, then the difference in length is one more reason to prefer RCd over Scale A.

RC1. Rouse et al. (2008) concluded that RC1 and its proxy, the HEA Content scale, measure the same attribute equally reliably and are essentially interchangeable. Identified empirical correlates for RC1 include various forms and indexes of somatoform psychopathology (Arbisi et al., 2008; Forbey & Ben-Porath, 2007b, 2008; Handel & Archer, 2008; Osberg et al., 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Tellegen et al., 2003), chronic pain and number of reported health complaints (Arbisi et al., 2008; Handel & Archer, 2008; Tellegen et al., 2003), and poor insight with respect to physical health (Wygant et al., 2007).

In regard to discriminant validity, depressed mood and other nonspecific distress symptoms in mental health settings are correlated with RC1 (as would be expected) but less so than with RCd (Arbisi et al., 2008; Forbey & Ben-Porath, 2008; Handel & Archer, 2008; Osberg et al., 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Tellegen et al., 2003). RC1 is generally uncorrelated with behavioral and thought dysfunction.

Similar correlates have been identified in the literature for HEA (although this literature is not as extensive), and the two scales have a similar somatic content. However, RC1 consists of fewer items: 27 versus 36. Of the two scales, HEA would be the one to be replaced, depending again on whether brevity is a consideration.

RC2. As we reported earlier, RCd and RC2 are more highly intercorrelated than the two designated proxies, A and INTR, respectively. Especially in Rouse et al.’s (2008) two clinical subsamples, the correlations between RCd and RC2 are substantial (.72 and .74), whereas the correlations between A and INTR are comparatively modest (.57 and .58). The relatively strong association between RCd and RC2 is clinically meaningful and called for given current understandings of Demoralization, namely, as both distinctive from and strongly related to core depression (see Tellegen et al., 2006). Its observed value conforms to Watson and Tellegen’s (1985) model of mood. The original version of the model entailed a .7 correlation of the pleasant-versus-unpleasant dimension (in other words, of Demoralization) with both (reversed) positive affect (measured by RC2) and with negative affect (measured by RC7). In the empirically updated hierarchical version of the model (Tellegen, Watson, & Clark, 1999) the latent correlation is somewhat higher still because positive affect and negative affect themselves are now recognized as modestly correlated.

The external correlates of RC2 are consistent with the nonendorsement of positive emotions reflected by high scores and corroborate its construct validity. RC2 is strongly correlated with various measures of (low) positive emotional temperament (PST) in broad-spectrum personality inventories—the Multi-dimensional Personality Questionnaire (MPQ; Sellbom & Ben-Porath, 2005), the NEO–PI–R (Sellbom et al., 2008b), and the SNAP (Simms & Clark, 2005). It is the strongest RC marker of the PEM construct on all three measures.

RC2 is correlated in particular with depressive mood symptoms (Arbisi et al., 2008; Forbey & Ben-Porath, 2007b; Handel & Archer, 2008; Osberg et al., 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Sellbom et al., 2008a; Tellegen et al., 2003, 2006) and is also a significant predictor of social anxiety (Forbey & Ben-Porath, 2008; Sellbom et al., 2008a). Specific depression markers, such as anhedonia and loss of interest, as well as nonspecific markers such as decreased appetite, sleep, concentration, energy, suicide, sense of worthlessness, and hopelessness (but no specific anxiety markers) are also associated with RC2 (Arbisi et al., 2008; Forbey & Ben-Porath, 2007a; Handel & Archer, 2008; Tellegen et al., 2003).

Supporting its discriminant validity are weak or substantially lower correlations of RC2 with negative temperament and anxiety disorders and essentially no association with externalizing or thought dysfunction (Arbisi et al., 2008; Forbey & Ben-Porath, 2008; Handel & Archer, 2008; Sellbom & Ben-Porath, 2005; Sellbom et al., 2008b; Simms & Clark, 2005).

In contrast to these results, Rouse et al.’s (2008) proposed proxy, INTR, although correlated (negatively) with some of the same positive temperament markers as RC2, assesses (as
intended) the broader introversion trait construct as is evident from the stronger correlations of INTR with therapist ratings of introversion (Sellbom, Ben-Porath, et al., 2006) and with the NEO-PI-R Extraversion domain scale (Bagby, Sellbom, Costa, & Widiger, 2008; Sellbom et al., 2008b) than were found for RC2.

Although RC2 and INTR are substantially correlated, the difference between INTR, a broad-gauged trait measure of personality/emotional temperament, and RC2, a more circumscribed affect-related scale, cannot be ignored and is no accident. Tellegen et al. (2003, pp. 21–22) reported in some detail the item selection and exclusion decisions that were made specifically to ensure not only the relatedness but also the distinctiveness of RC2 and social introversion. To seriously contemplate INTR as a replacement for RC2, a focused core indicator of mood disturbance, Rouse et al. (2008) must also have overlooked the obvious content differences between the two scales.

**RC3.** Rouse et al. (2008) reported that the proposed proxy of RC3, the CYN Content scale, boasts a higher mean IIC than RC3. Here and elsewhere, Rouse et al. have made a point of reporting effect sizes, presumably because these are more informative than the results of conventional null-hypothesis testing. However, Rouse et al. defeated the very purpose of doing so when they based final evaluations on small but “statistically significant” differences between effect sizes. Inspection of Rouse et al.’s (2008) Table 4 shows that the IICs of RC3 and CYN differ by .01 (.21 versus .20).

As we noted earlier, even if IIC or alpha differences are non-trivial, the higher values are not necessarily associated with more valid measurement. Of interest in this particular instance is Streiner’s (2003) observation that a high alpha may reflect unnecessary duplication of content. The CYN Content scale includes two items, 110 and 374, both stating that to benefit themselves, “most people will use somewhat unfair means,” whereas RC3 includes only Item 110. Too much content homogeneity will narrow the real-life relevance of a scale.

Rouse et al. (2008) also seemed to discount the misanthropic but “non-self-referential” content of RC3, which clearly distinguishes it from the persecutory and clearly “self-referential” content of RC6. Tellegen et al. (2003) noted that the two scales, used in combination, may tap into significant psychological differences. The CYN Content scale, on the other hand, mixes several clearly self-referential (although not frankly persecutory) statements with the majority of non-self-referential statements.

The empirical correlates of RC3 support its construct validity. High RC3 scores are associated with hostility, anger, and low trust (Sellbom, Ben-Porath, & Bagby, 2008b), negative beliefs about others (Forbey & Ben-Porath, 2007b, 2008; Handel & Archer, 2008), alienation, and blame externalization (Sellbom & Ben-Porath, 2005; Sellbom, Ben-Porath, Baum, Erez, & Gregory, 2008). In a sample of law enforcement candidates, elevated RC3 scores at time of hire predicted a variety of on-the-job problems including citizen complaints, rude behaviors, abuse of authority, externalizing blame, and uncooperativeness (Sellbom, Fischler, & Ben-Porath, 2007). No similar set of empirical correlates has been identified in the literature for CYN.

Demonstrating its discriminant validity, RC3 is weakly correlated with internalizing psychopathology and is not associated with self-referential persecutory ideation (Arbisi et al. 2008; Forbey & Ben-Porath, 2008; Sellbom et al., 2008; Tellegen et al., 2003). In contrast, CYN is significantly correlated with paranoia (Ben-Porath, McCully, & Almagor, 1993; Sellbom, Graham, & Schenk, 2005), more so than RC3 (Sellbom, Graham, et al., 2006), consistent with the more purely non-self-referential content of RC3.

**RC4.** Rouse et al. (2008) compared RC4, a broad measure of antisocial conduct and a history of juvenile conduct problems, with the AAS scale, which they described as a “face valid measure of symptoms of alcohol or drug problems” (p. 440). On the grounds that RC4 shares “the majority of its reliable variance” with the AAS scale, Rouse et al. (2008) suggested that it would be useful to “clarify whether [RC4] taps the broader construct of antisocial problems, as does its parent Scale 4, or if its correlates are more specific to substance abuse issues” (p. 440). Rouse et al. must have overlooked the wide-ranging antisocial content of RC4 as well as the extensive body of previously reported findings linking this scale not only to substance abuse but also to a broad range of behaviors and personality traits that are neither targeted nor assessed by AAS.

RC4 is correlated with a history of juvenile delinquency and adult criminal conduct—self-reported and recorded by others (Arbisi et al., 2008; Handel & Archer, 2008; Sellbom et al., 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Ben-Porath, & Stafford, 2007; Tellegen et al., 2003); alcohol and substance abuse, both current and lifetime (Arbisi et al., 2008; Forbey & Ben-Porath, 2007b, 2008; Sellbom et al., 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Ben-Porath, et al., 2007; Tellegen et al., 2003); aggressive and violent behavior, low received ratings of confidence in the ability to inhibit violence, and poor treatment outcome and recidivism in men with a history of committing domestic violence (Sellbom et al., 2008); family dysfunction, including a history of abusive and emotionally cold home environments (Forbey & Ben-Porath, 2007b; Wygant et al., 2007); increased risk for problematic conduct for law enforcement personnel (Sellbom, Fischler, et al., 2007); and poor adherence to surgery follow-up in a medical setting (Wygant et al., 2007).

In the personality domain, high RC4 scores are associated with low constraint, particularly with low behavioral control, low agreeableness and conscientiousness, impulsiveness and anger/aggression (Forbey & Ben-Porath, 2007a, 2007b, 2008; Sellbom & Ben-Porath, 2005; Sellbom et al., 2008b), and psychopathic personality traits—primarily those of the “impulsive-antisociality” factor (Benning, Patrick, Hicks, Blonigen, & Krueger, 2003; Sellbom, Ben-Porath, et al., 2007; Sellbom, Ben-Porath, Lilienfeld, Patrick, & Graham, 2005). In many of the studies just cited, the correlations of RC4 with a wide range of construct-relevant indicators substantially exceed those obtained with Clinical Scale 4.

**RC6.** As is evident from the content of RC6 and RC8, these two scales highlight the distinction between persecutory ideation and cognitive disorganization embedded in Clinical Scale 6 and 8, respectively. Rouse et al.’s (2008) choice of the PSY-5 PSYC scale and the BIZ Content scale as proxies for RC6 and RC8, respectively, does away with this critical difference. In contrast to RC6 and RC8, both PSYC and BIZ are broad spectrum and highly correlated measures of thought disorder. Both these proxies include most the RC6 items and RC8 items.
The “68” data point in Figures 1a through 1d brings out the marked difference in differentiation between the two RC scales and the two proxies. Another reason why the choice of PSYC as the RC6 proxy makes little sense clinically and conceptually is that PSYC was designed to measure a construct associated with Axis II personality pathology (primarily Cluster A) and not persecutory ideation (Harkness & McNulty, 1994).

The empirical correlates of RC6 show it to be a valid indicator of persecutory ideation and paranoid thinking. They include delusions, particularly of a persecutory nature, and ideas of reference (Arbisi et al., 2008; Handel & Archer, 2008; Tellegen et al., 2003), paranoia and interpersonal mistrust (Sellbom, Graham, et al., 2006; Simms & Clark, 2005), alienation and blame externalization (Handel & Archer, 2008; Sellbom & Ben-Porath, 2005; Sellbom et al., 2008), and the paranoia and mistrust aspects of Cluster A personality pathology (Simms & Clark, 2005).

As would be expected, RC6 is also correlated with a broader range of psychotic symptoms, including hallucinations and non-paranoid delusions, but the associations with the more specific criteria listed earlier are generally stronger. In further evidence of its discriminant validity, RC6 is generally uncorrelated with internalizing and behavioral dysfunction and with cynicism (Arbisi et al., 2008; Forbey & Ben-Porath, 2007a, 2008; Handel & Archer, 2008; Sellbom et al., 2008; Tellegen et al., 2003).

Given its conceptual basis, it is not surprising that PSYC is more highly correlated than RC6 with personality traits not associated with persecutory ideation (e.g., MPQ Absorption) (Harkness, McNulty, Ben-Porath, & Graham, 2002; Sellbom & Ben-Porath, 2005; Sellbom et al., 2008b). RC6, on the other hand, is more highly correlated than PSYC with variables measuring actual psychosis (Arbisi et al., 2008; Harkness et al., 2002; Tellegen et al., 2003) and persecutory delusions in particular.

**RC7.** We showed earlier that Rouse et al.’s (2008) proxies for RC2 and RC7 (our Demoralization measure), namely, INTR and Welsh’s Scale A, respectively, are too far apart. In startling contrast to their treatment of RC2, an index of (low) Positive Emotionality, Rouse et al. interpreted RC7, an index of Negative Emotionality, as indistinguishable from Demoralization. That is, Rouse et al. chose Scale A as a proxy for both RC7 and RC6.

The discrepancy between these proxy choices for the (RCd, RC2) and (RCd, RC7) pairs of scales is easily recognized in Figures 1a through 1d, which show data points d2 and d7 as widely separated in opposite directions from the diagonal.

As is true for RCd and RC2, RCd and RC7, although highly correlated (.78 and .79, respectively) in Rouse et al.’s (2008) two clinical subsamples, are not interchangeable. The difference in content between the two scales is also an indication, the RC7 items describing a more activated negative-affective tendency. As in the case of RC2, the correlational pattern observed for RC7 is consistent with Watson and Tellegen’s (1985) mood model, particularly Tellegen et al.’s (1999) updated version.

The empirical correlates of RCd and RC7 are likewise distinctive, contrary to Rouse et al.’s (2008) notion of interchangeability. The two scales clearly differentiate distress and fear symptomatology (e.g., Sellbom et al., 2008a; Tellegen et al., 2006), RC7 being more strongly associated with fear disorders than RCd.

More generally, RC7 is substantially correlated with measures of negative temperament as defined in major personality models, more strongly than with measures of positive temperament with which they tend to be weakly or uncorrelated (Sellbom & Ben-Porath, 2005; Sellbom et al., 2008b; Simms & Clark, 2005). In addition, RC7 is, of all the RC scales, the best predictor of fear-based disorders (e.g., panic, social phobia, specific phobia, agoraphobia) and obsessive–compulsive disorder symptoms (Sellbom et al., 2008a; Tellegen et al., 2006). RC7 is also the best predictor of anger and hostility (Forbey & Ben-Porath, 2007a; Sellbom et al., 2008b; Sellbom, Ben-Porath, et al., 2006).

With regard to discriminant validity, RC7 is generally weakly or hardly correlated with externalizing disorders and thought dysfunction (Arbisi et al., 2008; Forbey & Ben-Porath, 2007a, 2008; Handel & Archer, 2008; Sellbom et al., 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Tellegen et al., 2003), and as already mentioned, it is clearly less correlated with positive activation than with negative activation measures. In contrast, RC7’s proposed proxy, Scale A, is not only (like RC7) correlated with measures of anxiety but (like other first-factor markers) is more highly correlated than RC7 with depressed mood and general distress (e.g., Graham et al., 1999; Sellbom, Ben-Porath, et al., 2006).

**RC8.** As noted earlier, RC6 and RC8 represent major distinctive content and structural components of their respective parent Clinical scales. As a result, persecutory ideation and paranoid thinking are the primary and distinctive correlates of RC6 (noted earlier), whereas the strongest correlates of RC8 include auditory and visual hallucinations (Arbisi et al., 2008; Handel & Archer, 2008), nonpersecutory delusions, and being prescribed antipsychotic medication (Arbisi et al., 2008). In the area of personality and personality disorder variables, RC8 is saliently correlated with absorption, eccentric thinking, and schizotypal characteristics—including perceptual aberration and magical ideation (Forbey & Ben-Porath, 2007b, 2008; Sellbom & Ben-Porath, 2005; Simms & Clark, 2005). Indicative of its discriminant validity, the correlations of RC8 with persecutory delusions and with internalizing and behavioral dysfunction are lower (Arbisi et al., 2008; Forbey & Ben-Porath, 2007a, 2008; Handel & Archer, 2008; Sellbom et al., 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Tellegen et al., 2003).

In an outpatient mental health clinic sample, Rouse et al.’s (2008) proxy for RC8, BIZ, is correlated with psychotic symptoms, paranoid ideation, and hallucinations (Graham et al., 1999). However, BIZ adds only minimally to the contribution of RC8 as a predictor of psychotic symptoms, whereas RC8 adds significantly to BIZ (Sellbom, Graham, et al., 2006). RC8 not only surpasses BIZ in convergent validity, as noted here, but in view of its lower correlation with Paranoia/Mistrust, it bests BIZ also discriminantly (Sellbom, Graham, et al., 2006).

**RC9.** Rouse et al. (2008) identified the Ho scale as coming closest to qualifying as a proxy for RC9, but because the average correlation between the two scales is modest (.66), they concluded that RC9 and Ho “are measuring somewhat distinct constructs” (p. 439). Even a cursory inspection of the content of the two scales reveals why the constructs are more than somewhat distinct. In contrast to RC9, Ho includes a large cynicism component: It contains 11 of the 15 RC3 items (and 17 of the 23 CYN items). Cynicism is not part of RC9 because its parent...
scale, Clinical Scale 9, has no major cynicism component (and contains no RC 3 items).

The external correlates of RC 9 are consistent with its emotional and behavioral content. They include symptoms of manic episodes, such as experiencing racing thoughts, and increased likelihood of being prescribed mood stabilizers at discharge from a hospital (Arbisi et al., 2008). RC 9 is also substantially correlated with behavioral aggression (Forbey & Ben-Porath, 2007a). In a sample of domestic violence offenders, it was negatively correlated with received ratings of confidence in the ability to stop acting violently (Sellbom et al., 2008). Other correlates in this population include poorer treatment outcome and increased risk for recidivism (Sellbom et al., 2008).

Correlates of RC 9 in the personality and personality disorder domain include narcissism, manipulativeness, and dominance/social potency (Handel & Archer, 2008; Sellbom & Ben-Porath, 2005; Simms & Clark, 2005). The strongest personality correlates of RC 9 are a low level of FFM Agreeableness and (low) MPQ Harm-Avoidance (Sellbom & Ben-Porath, 2005; Sellbom et al., 2008b). RC 9 has also been found to be a marker of the psychopathy construct (Sellbom, Ben-Porath, et al., 2005; Sellbom, Ben-Porath, & Stafford, 2007). Other personality correlates of RC 9 include sensation seeking and/or behavioral fearlessness (Sellbom et al., 2008b; Sellbom, Ben-Porath, et al., 2005; Simms & Clark, 2005) and disinhibition and impulsivity (Forbey & Ben-Porath, 2008; Simms & Clark, 2005).

Demonstrating its discriminant validity, the correlations of RC 9 with internalizing psychopathology and thought dysfunction are generally weak or negligible (Arbisi et al., 2008; Forbey & Ben-Porath, 2007a, 2008; Handel & Archer, 2008; Sellbom et al., 2008; Sellbom, Ben-Porath, et al., 2006; Sellbom, Graham, et al., 2006; Tellegen et al., 2003). In contrast, Rouse et al.'s (2008) chosen proxy, Ho, is substantially less correlated than RC 9 with manic symptoms (Sellbom, Graham, et al., 2005, 2006) and more highly correlated with general maladjustment/distress, cynicism, and paranoia/mistrust (Graham et al., 1999; Sellbom, Graham, et al., 2005, 2006; Tellegen et al., 2003). Anger, aggression, and hostility are common correlates of RC 9 and Ho, but the correlations of Ho with these affectively negative manifestations are higher.

**SUMMARY**

Rouse et al. (2008) repeated the assertion that correlations of the RC scales with their parent Clinical scales are modest compared to the correlations with other existing MMPI–2 scales. In response, we reiterate a point we have stressed before: The RC scales were not intended to mimic the heterogeneous and overlapping content of the Clinical scales. Instead, the RC scales measure distinctive Clinical scale components that are individually focused but collectively wide ranging and allowing versatile predictive applications. Rouse et al. also revived the claim that most RC scales are redundant with existing MMPI–2 scales. Refuting this proposition once more, we have shown that the RC scale set accounts for Clinical scale variance more successfully and more economically than does Rouse et al.'s set of proxy scales. We note additional problems with Rouse et al.'s critique: problematic interpretations of reliability in relation to scale unidimensionality and to scale validity, failure to address correlational artifacts owing to scale overlap, and an additional repetition of a previously refuted criticism of the RC scales elevations.

At a fundamental level, Rouse et al. (2008) appeared not to recognize that to ask what a scale measures is to inquire into its construct validity. Rouse et al. not only restricted the scope of their internal analyses, they also neglected a large body of external correlates and relevant content characteristics essential to evaluating the construct validity of the RC scales. Addressing this deficiency, we have summarized, compared, and contrasted external correlates findings and pertinent content similarities and differences for each RC scale and its proposed proxy.

**REFERENCES**


Tellegen, A. (1982). Brief manual of the Multidimensional Personality Questionnaire. Unpublished manuscript, University of Minnesota, Minneapolis. Please verify University of Minnesota campus is Minneapolis, as inserted, for Tellegen, 1982, or correct.


