Detecting Negative Response Bias with the Fake Bad Scale, Response Bias Scale, and Henry–Heilbronner Index of the Minnesota Multiphasic Personality Inventory-2

Kimberly E. Dionysus 1,*, Robert L. Denney 1, Dale A. Halfaker 2

1The School of Professional Psychology at Forest Institute, Springfield, MO, USA
2Neuropsychological Associates of Southwest Missouri, Springfield, MO, USA

*Corresponding author at: 73 B Mountain Ave, Springfield, NJ 07081, USA. Tel.: +1-303-995-8170; fax: +1-973-313-1666.
E-mail address: cheyanne412@hotmail.com (K.E. Dionysus).

Accepted 11 November 2010

Abstract

The 28-item Response Bias Scale (RBS) and 15-item Henry–Heilbronner Index (HHI) are new validity scales within the Minnesota Multiphasic Personality Inventory-2, designed to detect over-reporting of cognitive and somatic symptomology, respectively. The 43-item Lees-Haley Fake Bad Scale (FBS) was designed to detect noncredible symptom presentations within a personal injury setting. The current study examined the predictive validity of these scales in a criterion-groups design involving head-injured litigants. Archival data were collected and two groups were created using the Slick, Sherman, and Iverson (1999) Criteria for Probable Negative Response Bias. Results yielded excellent to acceptable discrimination ability for the validity scales, including an area under the curve of 0.83 for FBS, 0.82 for RBS, and 0.73 for HHI. Findings suggest that the FBS, RBS, and HHI perform as well as the other validity scales in discriminating over-reporting within a head-injured litigant setting.

Keywords: Symptom validity testing; Traumatic brain injury; Forensic neuropsychology; Assessment

Introduction

Negative Response Bias (NRB) is a systematic tendency to produce more deficient scores than would be expected based on the skill level of the person (Franzen and Iverson, 2000). The Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) is often used in personal injury evaluations and has shown to be sensitive to NRB, particularly the Lees-Haley Fake Bad Scale (FBS; Larrabee, 2005). The Response Bias Scale (RBS; Gervais, Ben-Porath, Wygant, & Green, 2007) and Henry–Heilbronner Index (HHI; Henry, Heilbronner, Mittenberg, & Enders, 2006) have been introduced to enhance the evaluation of symptom validity. To date, the RBS and HHI have not been compared with the FBS in the context of a personal injury population. This study explores the diagnostic efficiency of the FBS, RBS, and HHI in predicting probable NRB, as defined by the Slick and colleagues (1999) Criteria.

Fake Bad Scale

The FBS consists of 43 items that were selected based on the premise that individuals feigning symptoms will attempt to appear as healthy, honest individuals who have experienced a difficult injury (Lees-Haley, English, & Glenn, 1991). In a review of the literature, the FBS was found to demonstrate sensitivity (SENS) to illogical symptom histories (Greiffenstein, Baker, Gola, Donders, & Miller, 2002), has been validated in detecting cognitive effort (Slick, Hopp, Strauss, & Spellacy, 1996),
and showed exceptional classification accuracies in the identification of mild head-injured litigants with poor cognitive effort (Ross, Millis, Krukowski, Putnam, & Adams, 2004).

Miller and Donders (2001) found that compensation-seeking individuals with mild traumatic brain injury (TBI) were twice as likely to elevate the FBS when compared with those who were not compensation seeking. Larrabee (2005) found the FBS to be more sensitive to exaggerated symptoms reported in civil litigation for neuropsychological claims, where neurocognitive dysfunction is more likely to be malingered. In fact, several studies have indicated that elevated FBS scores can occur with symptom exaggeration as well as malingered neurocognitive dysfunction (Larrabee, 1998, 2003a, 2003b; Slick et al., 1999). More recently, the literature surrounding the effectiveness of the FBS has been expanded upon to include an updated composite effect size of \(d = 0.95\), which is noted to be stable and relative to previous findings. In fact, this updated look into the FBS’ cumulative literature notes continuing group differentiating ability and a large effect size difference related to effort and TBI (Nelson, Hoelzle, Sweet, Arbisi, & Demakis, 2010).

A number of studies have identified different cut-off scores for the FBS that are effective in various settings. Originally, Lees-Haley and colleagues (1991) suggested a cut-off score of 20; this was later increased and modified for gender differences in a follow-up study (Lees-Haley, 1992). The more recent cut-off score recommended for women was \(\geq 26\), which resulted in an SENS of 74% and an associated specificity (SPEC) of 92%. The cut-off score recommend for men was \(\geq 24\), which resulted in an SENS of 75% and an SPEC of 96%. Numerous studies have identified a variety of cut-off scores, which vary by population. Larrabee (2003b) recommended other assessment tools be used in conjunction with FBS when FBS scores fall below 30.

A great deal of research identifies the FBS as an effective measure of NRB in personal injury populations; however, there are limitations. Rogers, Sewell, and Ustad (1995) caution that the SENS and SPEC of the FBS decreases in psychiatric settings, whereas many researchers have demonstrated that elevated FBS scores do not necessarily indicate malingered neurocognitive dysfunction. Confounding variables, such as intensity of head injury or preinjury psychiatric history, could increase the likelihood of false positives (Martens, Donders, & Millis, 2001). Butcher, Arbisi, Atlis, and McNulty (2003) suggested that the FBS demonstrated poor internal consistency, had an unacceptable high rate of false-positive identification in clinical groups, and over-identified malingered in a personal injury sample. In contrast, Lees-Haley and Fox (2004) noted, however, that the methodology used in the Butcher and colleagues research was unusual. Specifically, they argued that the FBS is intended to be used in a personal injury setting, whereas personal injury litigants constituted the smallest portion of the Butcher and colleagues (2003) samples. Moreover, the samples did not control for the effects of malingered.

The FBS has clearly demonstrated validity in the forensic setting (Nelson et al., 2010), but there is much less research on how the RBS and HHI perform in direct comparison to the FBS.

Response Bias Scale

The RBS is an empirically derived scale that consists of 28 MMPI-2 items, which predicted failure on the Computerized Assessment of Response Bias (Allen, Conder, Green, & Cox, 1997), the Word Memory Test (WMT; Green, 2005), and the Test of Memory Malingering (TOMM; Tombaugh, 1996), three commonly used Symptom Validity Tests (SVT). The scale is designed to be sensitive to cognitive effort that is associated with failure on cognitive SVT’s and is the only scale developed using SVT performance from a forensic disability clinical sample (Gervais et al., 2007). During initial validation, the authors examined the RBS within a sample of more than 300 participants and found that the scale yielded a large effect size \(d = 0.92\) in discriminating between individuals passing and failing the WMT and Medical Symptom Validity Test (Green, 2004a). This same study showed that the RBS added incrementally to the F, Fp, and FBS scales in predicting WMT performance. The RBS evidences ranges of SENS from 0.34, associated with a cutoff of 16, and 0.16 associated with a cutoff of 18. There was excellent SPEC, however (0.89 at a cutoff of 16 to 0.98 at a cutoff of 18). A higher cutoff of \(\geq 17\) is recommended to allow for a reduced risk of false positives, while achieving a high level of SPEC and increased positive predictive value.

The RBS has also evidenced discrimination of secondary gain within a criminal forensic sample (Nelson, Sweet, & Heilbronner, 2007). Similarly, Gervais, Ben-Porath, Wygant, and Green (2008) examined the relationship between the RBS, FBS, the MMPI-2 F-family, and self-reported memory complaints, as measured by the Memory Complaints Inventory (MCI; Green, 2004b). The RBS was significantly correlated with all MCI scales, more so than the MMPI-2 F-family and FBS. Gervais, Ben-Porath, Wygant, and Sellbom (2010) studied the RBS and the Restructured Form of the MMPI-2 (MMPI-2RF; Ben-Porath & Tellegen, 2008) and found that the RBS added incremental validity to the MMPI-2RF scales while exhibiting preferential ability to capture response bias.

Of importance to note are the conceptual differences between the RBS and the FBS. The FBS was derived using a combination of empirical and rational analyses, which aided in the selection of items that reflect exaggeration of post-injury emotional distress as well as the minimization of preinjury emotional problems (Greiffenstein, Fox, & Lees-Haley, 2007). Whitney, Davis, Shepard, and Herman (2008) found that in a sample of 46 VA Medical Center outpatients, the RBS was associated
with the largest effect sizes between groups who passed and failed the TOMM. They examined the incremental validity of the RBS using hierarchical regression analyses and concluded that the scale was superior to standard MMPI-2 validity scales as well as the HHI in predicting SVT failure.

More recently, the initial validation studies of the RBS have been expanded upon to examine the scale’s ability to predict SVT failure in a disability and criminal forensic sample. The disability sample results yield comparable performances to Gervais and colleagues’ (2007) results. Specifically, the RBS outperformed the F-Family and FBS in identifying SVT failure. In the criminal sample, the RBS demonstrated a large effect size in discriminating between valid and invalid defendant performance based on SVT’s (Wygant et al., 2010).

**Henry–Heilbronner Index**

The HHI consist of 15 items taken from a subset of the 17-item Pseudo-Neurologic Scale (PNS) used in the original MMPI (Hathaway & McKinley, 1943) and the FBS. It was developed to be sensitive to neurocognitive complaints expected in the months following head trauma. Henry and colleagues (2006) used a series of logistic regression analyses to determine whether the FBS, PNS, and a sum of the FBS and PNS (FBPNS) could differentiate between a probable malingering group and a nonmalingering group. The results demonstrated that all three were statistically significant predictors, although the FBS and FBPNS were slightly better. In order to distinguish 15 different items, correlations were obtained and those items with the strongest correlations (9 FBS items, 4 PNS items, and 2 shared items) were kept. Classification accuracies that maximized identification of malingering while minimizing false-positive errors were examined. A cutoff of ≥8 resulted in a classification accuracy rate of 85.6%, with a corresponding SPEC of 89% and an SENS of 80%. The results of this study suggest that the HHI may have utility in the evaluation of head-injured personal litigants (Henry, Heilbronner, Mittenberg, Enders, & Stanczak, 2008).

Exploring the diagnostic efficiency between the FBS, RBS, and HHI can provide useful information to clinicians working in litigation settings and potentially provide additional evidence that aids in the detection of NRB.

**Materials and Methods**

**Participants**

Participants were referred to a private practice for neuropsychological testing between 1999 and 2005. Patients were evaluated regarding a head injury in the context of litigation. Institutional Review Board approval was obtained and the study was in compliance with the ethical treatment of human participants. Archival data were extracted and de-identified. Participants were included if their files contained data for all 567 MMPI-2 items and at least three effort indicators. Using a criterion-groups design, participants were identified who appeared to exaggerate neurocognitive dysfunction. Criteria for inclusion and exclusion within noncredible and credible groups are described subsequently, including the use of performance measures of response bias for group assignment.

**Probable Negative Response Bias Group**

In total, 37 participants (age: \( M = 44.5, \ SD = 11.9 \); education: \( M = 12.1, \ SD = 1.8 \); Full-Scale Intelligence Quotient [FSIQ]: \( M = 88.2, \ SD = 9.0 \); 19 men and 30 women; 97.3% Caucasian and 2.7% African American) met the Slick and colleagues (1999) criteria for Probable Negative Response Bias (PNRB; Criterion B2). Specifically, all were in litigation or seeking to obtain disability benefits; all failed “one or more” well-validated psychometric tests or indices (tests and cutoffs are listed in Table 1) designed to measure exaggeration or fabrication of cognitive deficits not due to other psychiatric factors.

**Table 1. Effort indicators used for group assignments**

<table>
<thead>
<tr>
<th>Test/Indicator</th>
<th>Cut-Off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rey 15-Item Memory Test Recall Items (Lezak et al., 2004)</td>
<td>&lt;9</td>
</tr>
<tr>
<td>Finger Tapping Test (Larrabee, 2003c) total score</td>
<td>&lt;63</td>
</tr>
<tr>
<td>BCT Bolter VI infrequently missed items (Tenthula &amp; Sweet, 1996)</td>
<td>&gt;3</td>
</tr>
<tr>
<td>TOMM on either Trial 2 or retention (Tombaugh, 1997)</td>
<td>&lt;45</td>
</tr>
<tr>
<td>WMT failure on IR, DR, or CNS (Green, 2005)</td>
<td>&lt;82.5%</td>
</tr>
</tbody>
</table>

Notes: BCT = Booklet Categories Test; VI = validity index; TOMM = Test of Memory Malingering; WMT = Word Memory Test; IR = immediate recall; DR = delayed recall; CNS = test retest consistency.
neurologic, or developmental disorders. Although information regarding the severity of head injury was often not available, 25 participants reported post-traumatic amnesia congruent with mild head injury (< 1 h). The Glasgow ratings for eight participants were mild (at least 13 points) and one was moderate (9–12 points). The reported length of time the participants were unconscious due to their head injury indicates 17 to be in the mild range (length of unconsciousness \leq 30) and 5 to be in the moderate-to-severe range (length of unconsciousness > 30 min). Nineteen participants reported their head injuries to be due to motor vehicle accidents (51.4%), eight reported injuries caused by a fall (21.6%), nine reported injuries due to work related accidents (24.3%), and one reported injury due to an assault (2.7%). Sixteen participants were evaluated within 6 months to a year post-injury (43.2%), 12 were within 1- to 3-year post-injury (32.4%), five were within 3- to 5-year post-injury (13.5%), one was within 5- to 7-year post-injury (2.7%), two were within 7- to 10-year post-injury (5.4%), and one was tested more than 10-year post-injury (2.7%). In regards to psychiatric comorbidity, two participants were diagnosed with depression and three were diagnosed with an adjustment disorder.

Presumed Valid Group

The presumed valid (PV) group consists of 42 participants (age: \( M = 43.3, SD = 11.5 \); education: \( M = 12.2, SD = 2.1 \); FSIQ: \( M = 90.9, SD = 16.0 \); 25 men and 17 women; 95.2% Caucasian, 2.4% African American, and 2.4% Native American). All were involved in litigation; however, none of the participants had a failure on any effort measure. Although information regarding the severity of head injury was often not available, 22 participants reported post-traumatic amnesia congruent with mild head injury (< 1 h). The Glasgow ratings for 12 participants were mild (at least 13 points) and one was moderate (9–12 points). The reported length of time the participants were unconscious due to their head injury indicates 26 to be in the mild range (length of unconsciousness \leq 30 min). Twenty-six participants reported their head injuries to be due to motor vehicle accidents (61.9%), six reported injuries caused by a fall (14.3%), four participants reported injuries due to work-related accidents (9.5%), four reported their injury as other (9.5%), one reported injury due to an assault (2.4%), and one reported injury due to sports and recreation (2.4%). Nineteen participants received evaluations within 6 months to a year post-injury (45.2%), 18 were within 1- to 3-year post-injury (42.9%), three were within 3- to 5-year post-injury (71%), and two were within 7- to 10-year post-injury (4.8%). In regards to psychiatric comorbidity, one participant was diagnosed with a post-traumatic stress disorder and three with an adjustment disorder.

Procedures

MMPI-2 protocols were excluded if they had “Cannot Say” raw scores > 15 or elevations > 79T on VRIN or TRIN. The 28-item RBS and 15-item HHI raw scores were calculated for each case by the addition of items endorsed in the keyed direction, as suggested by Gervais and colleagues (2007) and Henry and colleagues (2006), respectively. Pearson’s product moment correlations were computed to determine whether any of the validity scales were significantly \( p \leq .05 \) associated with the FBS. To determine the predictive accuracy of the scales, a receiver-operating characteristic (ROC) curve was created and the diagnostic accuracy of the RBS and HHI was compared with the FBS. Various cut-off scores were evaluated by constructing cross-tabulation tables for the FBS, RBS, and HHI. Classification accuracies were evaluated by calculating the following: SENS, SPEC, positive predictive validity (PPV), and negative predictive validity (NPV). Cutoff scores were selected for the FBS, RBS, and HHI after analyzing the ROC curves and associated classification accuracy statistics.

Results

PNRB group failure rates on effort measures are detailed in Table 2. There were no failures on effort testing among the PV group. Preliminary analysis showed that the sample followed a normal test distribution. There was no significant group difference in age \( (p = .66) \), education \( (p = .77) \), ethnicity \( (p = .62) \), gender \( (p = .31) \), WAIS FSIQ \( (p = .36) \), or psychiatric comorbidity \( (p = .42) \). The PNRB group obtained statistically greater FBS scores \( (M = 25.35, SD = 4.44) \), RBS scores \( (M = 13.41, SD = 2.81) \) and HHI scores \( (M = 10.49, SD = 2.27) \) than the PV group. There was a significant difference between the PV and PNRB groups for FBS \( (p < .001) \), RBS \( (p < .001) \), and HHI \( (p < .001) \). Group statistics and effect sizes are outlined in Table 3.

Pearson’s product-moment correlation coefficients were computed for each of the predictor variables. Significant correlations were found between the raw scores of HHI and FBS \( (r = .73) \), HHI and RBS \( (r = .65) \), and FBS and RBS \( (r = .51) \).

An ROC curve was constructed for each scale. The ROC curve is a graphical plot of SENS versus 1-SPEC for a binary classifier system. The ROC curve has a diagonal line that represents the lower bound values for the curve, known as the
The ROC curve analysis produced excellent to acceptable area under the curve (AUC) for each scale (FBS \( \approx 0.83, \ SE = \pm 0.05; \) RBS \( \approx 0.82, \ SE = \pm 0.05; \) HHI \( \approx 0.73, \ SE = \pm 0.05). Each of the validity scales were above the chance diagonal line.

Logistic regression analyses were conducted to compare the RBS and HHI with the FBS in prediction of group membership. The results indicate that this sample was able to correctly classify 78.5% of PV and PNRB participants. Of the participants who were predicted to be PNRB, this model accurately picked 76.3% of them. Of the PV group, 80.5% were accurately predicted.

Elevated scores on the RBS (\( \chi^2 = 8.38, p = .004 \)) and FBS (\( \chi^2 = 7.09, p = .008 \)) significantly increased the model’s ability to predict NRB when compared with the HHI (\( \chi^2 = 0.02, p = .90 \)). Significance is obtained by values of \( p < .05 \). Refer to Table 4 for odds ratio and confidence intervals.

Optimal cutoff scores were selected from high levels of SPEC and the corresponding SENS. At roughly 90% SPEC, FBS has a cutoff of \( \geq 25 \) (SENS of 0.60; classification rate of 76%), RBS has a cutoff of \( \geq 14 \) (SENS of 0.58; classification rate of 68%), and HHI has a cutoff of \( \geq 12 \) (SENS of 0.45; classification rate of 69%). The PPV and NPV values corresponding to various cut-off scores for FBS, RBS, and HHI are described in Table 5.

### Discussion

The present study explores the diagnostic efficiency between the FBS, RBS, and HHI validity scales in detecting NRB in a head-injured litigant population. The FBS, RBS, and HHI were all significantly related and demonstrated excellent (FBS and RBS) to acceptable (HHI) discrimination ability. Earlier research found the FBS to be an effective measure of NRB in a personal injury population. The RBS has demonstrated an ability to outperform the F-family and FBS in identifying SVT failure (Wygant et al., 2010), whereas initial validation studies of the HHI suggest utility in a litigant population (Henry et al., 2008).

Significant relationships between these scales are not surprising, considering the amount of item overlap and their ability to measure cognitive bias. The HHI was derived from significant correlations within the FBS and FBPNS scales, thus sharing four items. The HHI is in essence, “FBS lite,” when the FBS goes up the HHI will also go up. The RBS and FBS also share four items, whereas the RBS and HHI share three items. The AUC of the FBS is 0.83 which is only slightly better than the AUC of RBS which is 0.82.
Due to the significant relationship between the three scales and the minimal difference in discrimination ability, further evaluation was conducted. This was done to evaluate what the RBS and HHI add to the FBS in predicting group membership. The results indicate that using the FBS ($\chi^2 = 7.09, p = .008$) and RBS ($\chi^2 = 8.38, p = .004$) significantly increased the model’s ability to detect NRB in this population sample better than when adding the HHI ($\chi^2 = 0.02, p = .90$). Both the FBS and RBS made significant contributions to the prediction of NRB. In fact, the predictive ability of the FBS increased with the addition of the RBS. As such, as scores are elevated on FBS and RBS, the likelihood of detecting NRB increases. In contrast, the HHI did not significantly contribute to group prediction and decreased the effectiveness of the FBS and RBS. The HHI appeared to provide no incremental validity to FBS and RBS.

Cutoff scores were also selected. Those selected varied from those obtained in the literature. The present study found an optimal cutoff score of $\geq 25$ for FBS. Lees-Haley and colleagues (1991) recommended FBS cutoff scores of $\geq 24$ for men and $\geq 26$ for women. These scores resulted in a classification rate of 75% for malingered men and 74% of malingered women. The present study revealed a slightly lower classification rate for the $\geq 24$ cutoff (74%) and a slightly higher rate (76%) for a cutoff of $\geq 26$. This variance is likely due to a smaller sample size and a lower ratio of men to women. Results pertaining to the RBS yielded a cutoff score of $\geq 14$, which is lower than the cutoff of $\geq 17$ (SPEC of 1.00, SENS of 0.14) as was originally proposed by Gervais and colleagues (2007). In addition, a cut-off score of $\geq 12$ was selected for the HHI. This score is greater than the score set forth by initial validations studies. Henry and colleagues (2006) suggested a cutoff of $\geq 8$. This cutoff yielded an SENS of 0.92 and an SPEC of 0.40.

### Table 4. Logistic regression results

<table>
<thead>
<tr>
<th>Scale</th>
<th>$B$ (SE)</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$p$-value</th>
<th>Exp($B$)</th>
<th>95% CI for Exp($B$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBS</td>
<td>0.24 (0.09)</td>
<td>7.09</td>
<td>1</td>
<td>.008</td>
<td>1.27</td>
<td>(1.06, 1.51)</td>
</tr>
<tr>
<td>RBS</td>
<td>0.40 (0.14)</td>
<td>8.38</td>
<td>1</td>
<td>.004</td>
<td>1.49</td>
<td>(1.14, 1.96)</td>
</tr>
<tr>
<td>HHI</td>
<td>0.02 (0.18)</td>
<td>0.02</td>
<td>1</td>
<td>.900</td>
<td>1.02</td>
<td>(0.71, 1.47)</td>
</tr>
</tbody>
</table>

*Notes:* FBS = Fake Bad Scale; RBS = Response Bias Scale; HHI = Henry–Heilbronner Index; $B$ = regression coefficient; SE = standard error; $\chi^2$ = Wald statistic; $p$ = significance; Exp($B$) = indicator in the change of odds; CI = confidence interval.

### Table 5. Classification accuracy statistics for FBS, RBS, and HHI in predicting PNRB at a 40% base rate

<table>
<thead>
<tr>
<th>Scale</th>
<th>Cutoff</th>
<th>SPEC</th>
<th>SENS</th>
<th>PPV</th>
<th>NPV</th>
<th>Class. rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>FBS</td>
<td>$\geq 21.00$</td>
<td>0.571</td>
<td>0.838</td>
<td>0.633</td>
<td>0.800</td>
<td>0.696</td>
</tr>
<tr>
<td></td>
<td>$\geq 22.00$</td>
<td>0.691</td>
<td>0.757</td>
<td>0.683</td>
<td>0.763</td>
<td>0.722</td>
</tr>
<tr>
<td></td>
<td>$\geq 23.00$</td>
<td>0.738</td>
<td>0.676</td>
<td>0.694</td>
<td>0.721</td>
<td>0.709</td>
</tr>
<tr>
<td></td>
<td>$\geq 24.00$</td>
<td>0.796</td>
<td>0.676</td>
<td>0.735</td>
<td>0.745</td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td>$\geq 25.00$</td>
<td>0.905</td>
<td>0.596</td>
<td>0.846</td>
<td>0.717</td>
<td>0.760</td>
</tr>
<tr>
<td></td>
<td>$\geq 26.00$</td>
<td>0.929</td>
<td>0.568</td>
<td>0.875</td>
<td>0.709</td>
<td>0.760</td>
</tr>
<tr>
<td></td>
<td>$\geq 27.00$</td>
<td>0.976</td>
<td>0.568</td>
<td>0.954</td>
<td>0.719</td>
<td>0.785</td>
</tr>
<tr>
<td></td>
<td>$\geq 28.00$</td>
<td>0.976</td>
<td>0.378</td>
<td>0.933</td>
<td>0.640</td>
<td>0.696</td>
</tr>
<tr>
<td></td>
<td>$\geq 29.00$</td>
<td>0.976</td>
<td>0.243</td>
<td>0.900</td>
<td>0.594</td>
<td>0.633</td>
</tr>
<tr>
<td>RBS</td>
<td>$\geq 13.00$</td>
<td>0.881</td>
<td>0.621</td>
<td>0.821</td>
<td>0.725</td>
<td>0.759</td>
</tr>
<tr>
<td></td>
<td>$\geq 14.00$</td>
<td>0.928</td>
<td>0.540</td>
<td>0.870</td>
<td>0.697</td>
<td>0.658</td>
</tr>
<tr>
<td></td>
<td>$\geq 15.00$</td>
<td>0.928</td>
<td>0.351</td>
<td>0.812</td>
<td>0.619</td>
<td>0.658</td>
</tr>
<tr>
<td></td>
<td>$\geq 16.00$</td>
<td>0.976</td>
<td>0.216</td>
<td>0.889</td>
<td>0.586</td>
<td>0.620</td>
</tr>
<tr>
<td></td>
<td>$\geq 17.00$</td>
<td>1.00</td>
<td>0.135</td>
<td>1.00</td>
<td>0.568</td>
<td>0.595</td>
</tr>
<tr>
<td>HHI</td>
<td>$\geq 7.00$</td>
<td>0.261</td>
<td>0.973</td>
<td>0.537</td>
<td>0.916</td>
<td>0.594</td>
</tr>
<tr>
<td></td>
<td>$\geq 8.00$</td>
<td>0.404</td>
<td>0.919</td>
<td>0.576</td>
<td>0.850</td>
<td>0.646</td>
</tr>
<tr>
<td></td>
<td>$\geq 9.00$</td>
<td>0.571</td>
<td>0.783</td>
<td>0.617</td>
<td>0.750</td>
<td>0.671</td>
</tr>
<tr>
<td></td>
<td>$\geq 10.00$</td>
<td>0.762</td>
<td>0.568</td>
<td>0.677</td>
<td>0.667</td>
<td>0.671</td>
</tr>
<tr>
<td></td>
<td>$\geq 11.00$</td>
<td>0.881</td>
<td>0.513</td>
<td>0.792</td>
<td>0.673</td>
<td>0.709</td>
</tr>
<tr>
<td></td>
<td>$\geq 12.00$</td>
<td>0.929</td>
<td>0.405</td>
<td>0.833</td>
<td>0.640</td>
<td>0.684</td>
</tr>
<tr>
<td></td>
<td>$\geq 13.00$</td>
<td>0.976</td>
<td>0.243</td>
<td>0.900</td>
<td>0.594</td>
<td>0.633</td>
</tr>
</tbody>
</table>

*Notes:* Cut-off scores are in raw score units. FBS = Fake Bad Scale; RBS = Response Bias Scale; HHI = Henry–Heilbronner Index; SPEC = specificity; SENS = sensitivity; PPV = positive predictive validity; NPV = negative predictive validity; Class. rate = classification rate.

Due to the significant relationship between the three scales and the minimal difference in discrimination ability, further evaluation was conducted. This was done to evaluate what the RBS and HHI add to the FBS in predicting group membership. The results indicate that using the FBS ($\chi^2 = 7.09, p = .008$) and RBS ($\chi^2 = 8.38, p = .004$) significantly increased the model’s ability to detect NRB in this population sample better than when adding the HHI ($\chi^2 = 0.02, p = .90$). Both the FBS and RBS made significant contributions to the prediction of NRB. In fact, the predictive ability of the FBS increased with the addition of the RBS. As such, as scores are elevated on FBS and RBS, the likelihood of detecting NRB increases. In contrast, the HHI did not significantly contribute to group prediction and decreased the effectiveness of the FBS and RBS. The HHI appeared to provide no incremental validity to FBS and RBS.
Limitations

This study has several limitations. The most significant is sample size. Although the sample has an $n = 79$, a larger sample size would have helped solidify the predictive power of the FBS, RBS, and HHI. Also, the entire population had some form of incentive to perform poorly in the form of pending litigation. This means that even those subjects who were classified in the PV group satisfied Slick and colleagues (1999) Criteria A. If this study were repeated, a cleaner, more specifically defined sample would increase homogeneity. Using the Slick and colleagues (1999) Criteria allows the researcher to use different effort measures; however, this can weaken the criteria. Additionally, using only Criterion B2 for classification of the two groups may have led to a less precise a priori classification.

Conflict of Interest

Dr. Denney serves on the editorial board of Archives of Clinical Neuropsychology; receives royalties from publication of Clinical Neuropsychology in the Criminal Forensic Setting (Guilford Press, 2008) and Ethical Practice in Forensic Psychology (American Psychological Association, 2006); and serves on the National Academy of Neuropsychology Board of Directors as Past President.

Acknowledgements

The authors would like to thank Roger Gervais, Ph.D., for his review and contributions to this work.

References


