PAPER

PSYCHIATRY & BEHAVIORAL SCIENCES

Glenn D. Walters,[1] Ph.D.

Screening for Malingering/Exaggeration of Psychiatric Symptomatology in Prison Inmates Using the PICTS Confusion and Infrequency Scales*

ABSTRACT: A sample of 76 federal prison inmates with a history or current complaints of significant psychiatric symptomatology at intake were followed for a period of 4–39 months by a psychologist who rated the inmate as malingering ($n = 12$), substantially exaggerating ($n = 32$), minimally exaggerating ($n = 23$), or honestly reporting ($n = 9$) signs and symptoms of schizophrenia, bipolar disorder, major depression, or severe anxiety disorder. The Confusion-revised (Cf-r) and Infrequency (INF) scales of the Psychological Inventory of Criminal Thinking Styles, which had been administered routinely at intake, revealed that only the INF successfully predicted malingering and exaggeration of psychiatric symptomatology even after pre-existing group differences in age, race, and overall criminal thinking were controlled. These results suggest that the INF scale can potentially serve as an effective initial screening measure for malingering/exaggeration in inmates presenting with mental health complaints.

KEYWORDS: forensic science, forensic assessment, malingering, PICTS, Infrequency scale

Identifying offenders who feign and exaggerate psychiatric symptomatology is an important topic in forensic psychology, and a variety of instruments and measures have been developed to assist psychologists in assessing, appraising, and screening for symptom feigning and exaggeration in forensic populations. Perhaps the most popular and extensively researched of these instruments is the Structured Interview of Reported Symptoms (SIRS) (1). Research indicates that the SIRS is a reliable and valid measure of dissimulation (2). However, it requires up to an hour to administer and is limited by the extensiveness of the examiner's observation of a restricted range of respondent behavior (3). In an effort to address the first concern, shorter interview schedules like the Miller–Forensic Assessment of Symptoms Test (M–FAST) (4), self-report measures like the Structured Inventory of Malingered Symptomatology (SIMS) (5), and validity scales embedded in multiscale inventories like the revised Minnesota Multiphasic Personality Inventory (MMPI-2) (6) and Personality Assessment Inventory (PAI) (7) have been devised. Preliminary research on the M-FAST, SIMS, Negative Impression scale and Malingering Index of the PAI, and F and Fp scales of the MMPI-2, often using the SIRS as the criterion measure of dissimulation, has generally supported the utility of these scales in screening for malingering and exaggeration of psychiatric symptomatology (3,8–11).

In addition to validity scales from general personality inventories like the MMPI-2 and PAI, there is at least one self-report measure specific to criminality that also contains embedded validity scales: i.e., the Psychological Inventory of Criminal Thinking Styles (PICTS) (12). The PICTS is an 80-item inventory designed to measure eight criminal thinking styles and several higher-order cognitive processes believed to be instrumental in advancing and maintaining a criminal lifestyle. Embedded in the PICTS are two validity or response style scales, Confusion and Defensiveness. The Confusion scale was created for the purpose of identifying reading problems, random responding, and "fake bad" response styles, whereas the Defensiveness scale was created to identify denial of common human foibles, defensive responding, and "fake good" response styles because both response styles can affect the elevation of the criminal thinking scales. When it was determined that the original PICTS validity scales suffered from low internal consistency and modest validity (13), they were subsequently revised. The revised scales achieved substantially better internal consistency and slightly better validity than the original scales, and while the revised Confusion scale (Cf-r) failed to distinguish between inmates instructed to simulate "fake good" and "fake bad" response styles, it did distinguish between inmates who had been instructed to simulate a "fake bad" response style and inmates who completed the PICTS under normal test-taking conditions (14).

Another response style scale was recently identified for the PICTS which may supplement or even improve upon the Cf-r as an index of malingering and exaggeration of psychiatric symptomatology. Initially identified in a factor analysis of the PICTS, the Infrequency scale (INF) was originally entitled Interpersonal Hostility because unlike the other three PICTS factors, there were no

[1]Federal Correctional Institution, Schuylkill – Psychology, PO Box 700, U.S. 81 & Route 901, Minersville, PA.

*The assertions and opinions contained herein are the private views of the author and should not be construed as official or as reflecting the views of the Federal Bureau of Prisons or the United States Department of Justice.

specific thinking style scales that loaded reliably on this factor and the content of several of the items implied the presence of social hostility (12). Just recently, it was discovered that the Interpersonal Hostility factor is actually a response style factor comprised of 10 items with some of the lowest rates of endorsement in the PICTS normative sample (10 INF items, range = 1.17–1.34, M = 1.23, SD = 0.39; 70 other PICTS items, range = 1.28–3.44, M = 1.90, SD = 0.38) (15). To the extent that both the MMPI-2 and PAI have scales formed from infrequently endorsed items, it was reasoned that the PICTS might also benefit from a validity indicator composed of items that are rarely endorsed by inmates. Reanalyzing the simulation data from the Walters' (14) investigation, it was discerned that scores on the INF scale were significantly higher in the group of inmates instructed to adopt a "fake bad" response style than in the group of inmates completing the PICTS under standard test-taking conditions and inmates simulating a "fake good" response style and in so doing outperformed the Cf-r scale. The next question that needs to be answered is whether the PICTS Cf-r and INF scales are capable of identifying malingered and exaggerated reports of psychiatric symptomatology in known groups of inmates, some of whom are identified as malingering, others of whom are identified as exaggerating, and still others who are identified as genuinely disturbed.

The purpose of this study was to ascertain whether the PICTS Cf-r and INF scales can serve as effective screening measures for malingered and exaggerated psychiatric symptomatology in a group of federal prison inmates. In this study, malingering and exaggeration were assessed using a graduated four-level rating scheme: level 1, malingering with no evidence of serious mental health problems; level 2, significant exaggeration in an individual with significant mental health problems; level 3, mild exaggeration in an individual with significant mental health problems; and level 4, significant mental health problems with no evidence of exaggeration. It was predicted that the Cf-r and INF scales would correlate significantly with dichotomized (malingering/nonmalingering) and quasi-continuous (four levels of gradually increasing exaggeration) measures of outcome, that both scales would achieve significant receiver operating characteristic (ROC) coefficients, and that both scales would record reasonable sensitivity and specificity at three standard cut scores (T-score = 60, T-score = 70, and T-score = 80). It was further reasoned that Cf-r and INF would achieve significant predictive effects when each was forced into a logistic regression (dichotomized outcome) and linear regression (quasi-continuous outcome) multivariate equation behind demographic and general criminal thinking variables that correlate univariately with the dichotomized and quasi-continuous outcomes.

Method

Participants

Participants for this study were 76 male inmates admitted to a medium-security federal correctional facility located in the northeast region of the United States between July 2003 and March 2007. During the intake interview, each participant reported or displayed (through a records review) signs and symptoms of serious emotional disorder. The average age of these 76 inmates was 36.68 years (SD = 9.24) and the mean educational level was 11.57 years (SD = 2.08). Ethnically, 56.6% of the sample was black, 35.5% was white, 6.6% was Hispanic, and 1.3% was Asian/Native American. Nearly three-quarters of the sample listed their marital status as single (71.1%), with 10.5% describing themselves as married,

17.1% describing themselves as divorced, and 1.3% describing themselves as widowed. The modal confining offense was a drug offense (28.9%), followed by a firearms violation (25.0%), a violent crime (21.1%), robbery (17.1%), and a property offense (1.3%); the remainder of the sample (6.6%) was convicted of miscellaneous crimes such as fraud and counterfeiting.

Measure

The PICTS (12) is an 80-item self-report measure designed to assess eight thinking styles held to support, protect, and reinforce a criminal lifestyle. Each item consists of a statement that is rated on a four-point Likert-type scale (strongly agree = 4 points, agree = 3 points, uncertain = 2 points, and disagree = 1 point). The standard PICTS protocol yields two 8-item validity scales—Confusion-revised (Cf-r) and Defensiveness-revised (Df-r)—eight 8-item thinking style scales—Mollification (Mo), Cutoff (Co), Entitlement (En), Power Orientation (Po), Sentimentality (Sn), Superoptimism (So), Cognitive Indolence (Ci), and Discontinuity (Ds)—four 10-item factor scales—Problem Avoidance (PRB), INF, Self-Assertion/Deception (AST), and Denial of Harm (DNH)—two content scales—Current (CUR) and Historical (HIS)—two composite scales—Proactive Criminal Thinking (P) and Reactive Criminal Thinking (R)—and a 64-item General Criminal Thinking (GCT) score. The two PICTS scales that served as predictors in this study, Cf-r and INF, possess reasonably good test–retest reliability after two, r (Cf-r) = 0.91 and r (INF) = 0.88, and 12, r (Cf-r) = 0.64–0.87 and r (INF) = 0.48–0.67, weeks (15).

Procedure

Individuals who participated in this study were inmates who either reported the presence or displayed a history of serious mental health problems (as indicated by a brief review of the psychology file) during the intake interview. The definition of serious mental health problems adopted in the present investigation entailed signs and symptoms of schizophrenia, other psychotic conditions (e.g., atypical psychosis, delusional disorder, schizoaffective disorder), bipolar disorder, major depression, or a severe anxiety disorder (e.g., obsessive-compulsive disorder, post-traumatic stress disorder). Individuals satisfying this definition were followed for a period of 4–39 months by the psychologist (GDW) who conducted the intake interview. From the start of the project in July 2003, the author (GDW) was essentially the only psychologist at the institution where this study took place and as such conducted all of the intake interviews. Since July 2003, all inmates admitted to this facility were administered the PICTS by a psychology technician. Accordingly, inmates with mental health problems were not singled out or afforded special attention; instead, they routinely completed the PICTS like all newly arrived inmates. Out of a total intake sample of 2055 new inmates, 76 (3.7%) displayed or reported signs and symptoms of serious mental health disorder at the time of intake. All 76 participants produced valid PICTS protocols using the following standard criteria: Cf-r ≤ 100T; missing items ≤10; item response variability >0 (i.e., not all "2" or "uncertain" responses). Scales on valid protocols with missing items were prorated using procedures outlined in the PICTS test manual (15).

The psychologist providing the ratings for this study had 25 years of correctional experience and observed each inmate at least twice a month for a period of 4–39 months. Observations occurred in a variety of settings (living unit, work, recreation, dining hall) and were supplemented by interviews with staff who

had daily contact with the inmate. Each rating was based on multiple observations gathered over time and across situations that were used to identify inconsistencies between reported symptoms and actual behavior, unrealistic symptoms, incompatible symptoms, presence or absence of a psychiatric history, and response to psychotropic medication. Participants were assigned to one of four levels based on a classification scheme originally designed and developed by Walters, White, and Greene (16): level 1 (L-1), malingering with no mental health problem; level 2 (L-2), a mental health problem with significant exaggeration of symptomatology; level 3 (L-3), a mental health problem with mild exaggeration of symptomatology; and level 4 (L-4), a significant mental health problem with no exaggeration of symptomatology. Levels were coded in an opposite direction (e.g., L-1 = 4, L-4 = 1) so that a high score on Cf-r or INF correlated positively with malingering/exaggeration. Three outcomes were created: a dichotomized malingering outcome (L-1 = 2; L-2/L-3/L-4 = 1), a quasi-continuous exaggeration outcome (L-1 = 4, L-2 = 3, L-3 = 2, L-4 = 1), and a dichotomized extreme group outcome (L-1 = 2; L-4 = 1).

In an attempt to minimize criterion contamination, the psychologist making the ratings did his best to avoid knowing the PICTS scores of inmates presenting with mental health difficulties. Nonetheless, 11 inmates from the present sample had been participants in an earlier PICTS research project and four other inmates had their PICTS results previously reviewed for clinical decision-making purposes. The psychologist who furnished the ratings for this study was consequently exposed to the PICTS results of 15 current participants prior to rating them on the four-level malingering/exaggeration scale. Therefore, despite a concerted effort to keep the predictor (PICTS) and criterion (ratings) variables separate, it is possible that some of the ratings were influenced by the psychologist's prior knowledge of the inmate's PICTS results. It should be noted, however, that the INF scale (formally known as the Interpersonal Hostility scale) was never included in any of these studies and, in fact, was not even considered a response style measure until just recently. To gauge the reliability of the ratings used in this study, a paraprofessional familiar with many of the participants in this study and who had no access to the PICTS results provided ratings independent of the psychologist for the 47 inmates (62% of the total sample) with whom he was familiar at the end of the rating period, the results of which revealed moderately good reliability on the four-level rating scale: Intraclass Correlation Coefficient (absolute agreement) = 0.75.

Results

Correlations

Twelve of the 76 inmates in this study (15.8%) were rated by the psychologist as malingering (L-1), 23 (30.3%) were rated as substantially exaggerating their symptoms (L-2), 32 (42.1%) were rated as minimally exaggerating their symptoms (L-3), and 9 (11.8%) were rated as honestly reporting symptoms of serious emotional disorder (L-4). Basic demographic measures (age, education, race, marital status, offense) and PICTS validity (Cf-r: M = 16.45, SD = 5.50; INF: M = 16.05, SD = 5.61) and criminal thinking (GCT: M = 130.71, SD = 34.98) scores were correlated with the three outcome measures: dichotomized malingering outcome, quasi-continuous exaggeration outcome, and dichotomized extreme group outcome. The resulting correlations are reproduced in Table 1 and indicate that age, race, INF, and GCT predicted all three outcomes,

while education, marital status, current offense, and Cf-r failed to predict a single outcome.

Classification Accuracy

ROC analysis of the dichotomized malingering outcome disclosed a lack of classification accuracy for Cf-r (area under the curve [AUC] = 0.660, $p = 0.08$, CI = 0.475–0.844) and significant effects for INF (AUC = 0.770, $p < 0.01$, CI = 0.583–0.957) and GCT (AUC = 0.725, $p < 0.05$, CI = 0.554–0.903). ROC analysis of the dichotomized extreme group outcome revealed that while Cf-r (AUC = 0.718, $p = 0.10$, CI = 0.489–0.946) failed to achieve an effect, INF (AUC = 0.810, $p < 0.05$, CI = 0.616–1.00) and GCT (AUC = 0.787, $p < 0.05$, CI = 0.590–0.984) successfully predicted the extreme group outcome.

For a scale, test, or instrument to be an effective screening device, sensitivity should be at least 0.80 and specificity should exceed 0.50 (17). As indicated by the results outlined in Table 2,

TABLE 1—Demographic and PICTS correlates of outcome.

	Outcome		
Variable	Malingering (n = 76)	Exaggeration (N = 76)	Extreme Group (n = 21)
Age	−0.34**	−0.47***	−0.77***
Education	0.04	0.06	0.32
Race	0.25*	0.26*	0.52*
Marital status	−0.04	−0.06	−0.20
Confining offense	0.12	0.08	0.09
Cf-r	0.22	0.19	0.34
INF	0.42***	0.41***	0.54*
GCT	0.29*	0.36**	0.44*

Coefficients for Age, Education, Df-r, Cf-r, INF, and GCT are Pearson product moment and point-biserial correlations and coefficients for Race, Marital Status, and Confining Offense are point-biserial and phi correlations.

Malingering, dichotomized malingering outcome; Exaggeration, quasi-continuous exaggeration outcome; Extreme Group, dichotomized extreme group outcome; Age, chronological age in years; Education, educational level in years; Race, white (1) versus nonwhite (2); Marital status, single (1) versus nonsingle (2); Confining offense, person crime (1) versus nonperson crime (2); Cf-r, PICTS Confusion-revised scale; INF, PICTS Infrequency scale; GCT, PICTS General Criminal Thinking score; PICTS, Psychological Inventory of Criminal Thinking Styles.

*$p < 0.05$.
**$p < 0.01$.
***$p < 0.001$.

TABLE 2—Classification accuracy of the INF and Cf-r scales in predicting the dichotomized malingering outcome at three different cut scores.

Scale	Cut Score	Hit Rate	SEN	SPEC	PPP	NPP
INF	≥60T	0.68	0.83	0.66	0.31	0.95
	≥70T	0.76	0.67	0.78	0.36	0.93
	≥80T	0.83	0.58	0.88	0.47	0.92
Cf-r	≥60T	0.45	0.75	0.39	0.19	0.89
	≥70T	0.63	0.67	0.62	0.25	0.91
	≥80T	0.80	0.50	0.86	0.40	0.90

INF, Psychological Inventory of Criminal Thinking Styles (PICTS) Infrequency scale; Cf-r, PICTS Confusion-revised scale; Cut Score, T-score at which respondent is identified as malingering; Hit Rate, overall classification accuracy; SEN, sensitivity (true positives/true positives + false negatives); SPEC, specificity (true negatives/true negatives + false positives); PPP, positive predictive power (true positives/true positives + false positives); NPP, negative predictive power (true negatives/true negatives + false negatives); N = 76.

the INF scale achieved 0.83 sensitivity and 0.66 specificity at a cut score of 60T. Higher cut scores attained more accurate results but at the expense of a rising false negative rate and associated drop in sensitivity—hence making these cut scores less practical for screening purposes. The Cf-r scale, by comparison, failed to obtain a sensitivity of ≥0.80 at any cut score.

Incremental Validity

The incremental validity of the INF scale was tested by forcing it into regression equations behind the other three variables that achieved univariate significance in this study: i.e., age, race, and GCT. A logistic regression analysis of the dichotomized malingering outcome in which age, race (white–nonwhite), and GCT (minus the eight items found on the INF) were entered at the first block of the equation and INF was entered at the second block disclosed significant effects for both age ($b = -0.150$, SE = 0.065, Wald = 5.39, $p < 0.05$) and INF ($b = 0.261$, SE = 0.109, Wald = 5.70, $p < 0.05$). Likewise, a linear regression analysis of the quasi-continuous exaggeration outcome measure where age, race (white–nonwhite), and GCT (minus items found on INF) were entered at the first step of the equation and INF was entered at the second step also identified significant effects for age ($b = -0.038$, SE = 0.010, $\beta = -0.385$, $t = -3.74$, $p < 0.001$) and INF ($b = 0.050$, SE = 0.021, $\beta = 0.309$, $t = 2.31$, $p < 0.05$). When the Cf-r was inserted into the second block/step of a logistic or linear regression analysis, the results were nonsignificant ($p > 0.10$).

Supplemental Analyses

The 15 PICTS protocols that the psychologist providing the ratings had foreknowledge of were removed from the sample, and the analyses recalculated to determine whether this foreknowledge was responsible for the positive results recorded by the INF scale. In the new sample ($n = 61$), the INF scale correlated 0.46 ($p < 0.001$) with the dichotomized malingering outcome, 0.43 ($p < 0.001$) with the quasi-continuous exaggeration outcome, and 0.65 ($p < 0.01$) with dichotomized extreme group outcome. AUC values for the INF in this new sample were 0.827 ($p < 0.01$, CI = 0.594–1.06) as a predictor of the dichotomized malingering outcome and 0.866 ($p < 0.05$, CI = 0.646–1.09) as a predictor of the dichotomized extreme group outcome. The INF achieved a hit rate of 0.66, sensitivity of 0.86, and specificity of 0.63 at a cut score of 60T in the new sample. Finally, incremental validity was of borderline significance when the INF scale was loaded into a logistic regression equation behind age, race, and GCT in an attempt to predict the dichotomized malingering outcome ($b = 0.242$, SE = 0.129, Wald = 3.50, $p = 0.06$) and was statistically significant when the INF was loaded into a linear regression equation behind age, race, and GCT in an attempt to predict the quasi-continuous exaggeration outcome ($b = 0.047$, SE = 0.023, $\beta = 0.317$, $t = 2.06$, $p < 0.05$).

Discussion

As interest in forensic assessment has grown, so too has the demand for ways to identify defendants and inmates who feign and exaggerate psychiatric symptomatology. The present study sought to determine whether malingering/exaggeration could be predicted by two PICTS scales, Cf-r and INF, in a group of male prison inmates with a rate of malingering (15.8%) comparable to the rate reported in the literature (10–25%) (9,11,18,19). The Cf-r and INF scales were subjected to three tests: (i) correlations with three different outcomes (dichotomized malingering outcome,

quasi-continuous exaggeration outcome, and dichotomized extreme group outcome); (ii) ROC analyses of two different outcomes (dichotomized malingering outcome and dichotomized extreme group outcome); (iii) incremental validity analyses controlling for age, race, and general criminal thinking on two different outcomes (dichotomized malingering outcome and quasi-continuous exaggeration outcome). The results indicated that while the INF scale passed all three tests, the Cf-r scale failed every part of every test. It is not surprising that the INF scale performed as well as it did in the present study in that it adopts the highly regarded and well-validated rare symptom approach to malingering detection (2). The Cf-r, on the other hand, did not fare well in this study. Consequently, despite its adequate performance in screening for malingering/exaggeration in an earlier simulation study (14) and its moderately strong correlation with INF in the present study ($r = 0.71$), Cf-r failed to predict malingering/exaggeration of psychiatric symptomatology when a known groups design was employed.

The results of the present investigation suggest that INF may be an effective means of identifying inmates who present with mental health symptomatology that is feigned or exaggerated. However, for the INF to serve as a screening measure, it must possess certain features and characteristics. Of greater importance than a measure's high hit rate is its ability to identify the construct of interest, which in the present study was malingering/exaggeration. Bagby et al. (17) operationalize the criterion of identifying the construct of interest as a sensitivity value of 0.80 or higher. Additionally, the measure's specificity should be above 0.50 to avoid an excess number of false-positive determinations. Although false positives are less problematic for a screening device than are false negatives, it is important to avoid burdening the second-stage measure with an inordinate number of evaluations. Once screened, identified inmates can be assessed with a more comprehensive instrument, like the SIRS, which, while too expensive to administer to the entire inmate population, is an effective procedure for identifying the false positives that have made it through the first stage of the screening process. As indicated by the sensitivity and specificity data provided in Table 2, the INF displays the principal characteristics of a good screening measure at a cut score of 60. If, on the other hand, the goal is to use the INF to identify invalid PICTS protocols prior to conducting a clinical interpretation of the PICTS, then a cut score of 80 would be more appropriate given the fact that false-positive determinations (discarding a "valid" profile as invalid) are more problematic and costly than false-negative determinations when assessing the "validity" of a PICTS protocol.

The present study yielded three unanticipated outcomes that warrant further discussion. First, the rate of serious psychiatric disorder observed in this study was significantly lower than the rate normally reported for psychiatric disorder in correctional populations. In a meta-analysis of 62 surveys encompassing 18,530 male prisoners, Fazel and Danesh (20) obtained a 3.7% rate of psychosis and 10% rate of major depression. The 13.7% average rate of serious psychiatric disorder observed in the meta-analysis is four times higher than the 3.7% rate recorded in the present study and may reflect procedural (i.e., the use of more stringent diagnostic criteria) or setting (medium-security federal prison) differences between the current study and many of the studies included in the Fazel and Danesh (20) meta-analysis. Second, nonwhite inmates were disproportionally rated as symptom feigning or exaggerating relative to white inmates. Although this could potentially indicate a bias in the construction of criterion diagnoses, this same effect was observed in an independent sample of federal prisoners screened for malingering with the MMPI-2 and SIMS (9). Third, younger inmates

participating in the current study were more often evaluated as feigning or exaggerating psychiatric symptoms than older inmates. This finding is inconsistent with a recent study conducted on federal prison inmates in which malingering indicators on the PAI failed to correlate with age (21). There is insufficient evidence on the age and ethnic status correlates of malingering to make much of these findings, although the low prevalence rates of psychiatric disorder in the current study suggest that the sample used in the current study may have been somewhat atypical and that additional research is required using a larger sample of inmates housed in both state and federal facilities.

The strength of this study is that it made use of a clinically relevant and externally valid criterion in assessing malingering/exaggeration of psychiatric symptomatology. Instead of asking inmates to simulate a particular response style, a clinically relevant known groups design was employed, and instead of relying on a well-respected but fallible "gold standard" like the SIRS, the criterion measure was externally valid ratings made by an experienced clinician who knew and observed each participant for a period of 4–39 months. The cross-situational, cross-temporal, and internal diagnostic inconsistencies that furnish clues as to whether an inmate is malingering or exaggerating psychiatric symptomatology can be more readily discerned when the inmate is followed for a period of time and his or her behavior is observed by several different staff members than it can be from a single administration of a structured interview. Edens et al. (3) used both simulators and malingering diagnoses provided by a unit psychiatrist as their criterion and discovered that measures like the SIRS and SIMS identified the simulators but failed to identify clinically defined malingering. However, whereas malingerers in the Edens et al.'s (3) study were identified before being tested, participants in the present investigation were tested routinely at intake and only later identified as malingering or exaggerating psychiatric symptomatology.

Ironically, the principal strength of this study is also its greatest weakness. The clinical nature of the criterion measure enhances the external validity of the present study but brings into question its internal validity. Whereas every effort was made to keep knowledge of the intake PICTS separate from the subsequent ratings, 11 of the 76 inmates in the present study had participated in an earlier PICTS investigation and four other inmates had their PICTS results reviewed for clinical purposes. Despite attempts to base the psychologist's ratings exclusively on observations of the inmate and not on prior knowledge of the PICTS, the PICTS results of 15 participants were potentially available to the rating psychologist before he made his criterion ratings, thus raising the possibility of criterion contamination. Three factors, however, suggest that criterion contamination does not explain the present findings. First, when a paraprofessional blind to the PICTS and independent of the psychologist rated many of these same inmates, inter-rater reliability was moderately high. Second, the results of the INF scale were never included in any prior research and were never considered for clinical interpretive purposes because the scale's role as a response style measure was not recognized until just recently. Third, analyses conducted on the 61 PICTS to which the psychologist did not have access prior to making his criterion ratings did not significantly alter the INF's ability to predict malingering/exaggeration. Nevertheless, the small sample size and fact that the PICTS was administered and the malingering ratings made 4–39 months apart leave the present study open to a number of alternative interpretations.

The limitations of this study notwithstanding, the current results suggest several possible avenues for future research. First, the current study needs to be replicated in a larger sample of participants where predictor-criterion independence is assured. A larger sample would also make it possible to conduct separate analyses on the basis of an inmate's alleged motivation for malingering (e.g., transfer, hospitalization, amusement) (22). Second, there is a need to understand the latent structure of malingering and exaggeration of psychiatric symptomatology in prison and forensic populations. Many staff are of the opinion that there are two distinct classes of inmate who report psychiatric symptomatology: those who suffer from significant mental health problems and those who fake significant mental health problems. Research indicates, however, that feigned mental health symptomatology in forensic populations is a dimensional (degree of exaggeration) rather than categorical (malingers vs. genuinely disturbed) construct (23). Hence, there may be a large reservoir of individuals with significant mental health problems who periodically exaggerate these problems for secondary gain or in response to various situational demands and pressures. Examinees who exaggerate tend to cluster in levels 2 and 3 of Walters et al.'s (16) four-level scheme, the two largest groups of individuals in the current investigation. Therefore, even though the PICTS was created to assess criminal thinking patterns and its validity scales were designed to identify response styles that might artificially elevate or lower scores on the criminal thinking scales, validity scales like the INF which adopt a rare symptoms approach to detecting malingering may be of some use in screening for malingering/exaggeration in inmates who present with psychiatric symptomatology during intake.

References

1. Rogers R, Bagby RM, Dickens SE. Structured Interview of Reported Symptoms (SIRS) and professional manual. Odessa, FL: Psychological Assessment Resources, 1992.
2. Rogers R. Detection strategies for malingering and defensiveness. In: Rogers R, editor. Clinical assessment of malingering and deception, 3rd edn. New York, NY: Guilford, 2008;14–38.
3. Edens JF, Poythress NG, Watkins-Clay MM. Detection of malingering in psychiatric unit and general population prison inmates: a comparison of the PAI, SIMS, and SIRS. J Pers Assess 2007;88:33–42.
4. Miller HA. Miller–Forensic Assessment of Symptoms Test (M–FAST): professional manual. Odessa, FL: Psychological Assessment Resources, 2001.
5. Smith GP, Burger GO. Detection of malingering: validation of the Structured Inventory of Malingered Symptomatology. J Am Acad Psychiatry Law 1997;25:183–9.
6. Butcher JN, Dahlstrom WG, Graham JR, Tellegen A, Kaemmer B. Minnesota Multiphasic Personality Inventory—2: manual for administration and scoring. Minneapolis, MN: University of Minnesota Press, 1989.
7. Morey LC. The Personality Assessment Inventory (PAI): professional manual, 2nd edn. Lutz, FL: Psychological Assessment Resources, 2007.
8. Boccaccini MT, Murrie DC, Duncan SA. Screening for malingering in a criminal-forensic sample with the Personality Assessment Inventory. Psychol Assess 2006;18:415–23.
9. Lewis JL, Simcox AM, Berry DTR. Screening for feigned psychiatric symptoms in a forensic sample by using the MMPI–2 and the Structured Inventory of Malingered Symptomatology. Psychol Assess 2002;14:170–6.
10. Miller HA. The Miller-Forensic Assessment of Symptoms Test (M–Fast): test generalizability and utility across race, literacy, and clinical opinion. Crim Justice Behav 2005;32:591–611.
11. Rogers R, Ustad KL, Salekin RT. Convergent validity of the Personality Assessment Inventory: a study of emergency referrals in a correctional setting. Assessment 1998;5:3–12.

12. Walters GD. The Psychological Inventory of Criminal Thinking Styles: part I. Reliability and preliminary validity. Crim Justice Behav 1995;22: 307–25.
13. Walters GD. The Psychological Inventory of Criminal Thinking Styles: part II. Identifying simulated response sets. Crim Justice Behav 1995;22:437–55.
14. Walters GD. Revised validity scales for the Psychological Inventory of Criminal Thinking Styles (PICTS). J Offender Rehabil 2001;32:1–13.
15. Walters GD. Psychological Inventory of Criminal Thinking Styles (PICTS): professional manual. Allentown, PA: Center for Lifestyle Studies (Morrisville, NC: Lulu), 2006.
16. Walters GD, White TW, Greene RL. Use of the MMPI to identify malingering and exaggeration of psychiatric symptomatology in male prison inmates. J Consult Clin Psychol 1988;56:111–7.
17. Bagby RM, Rogers R, Buis T. Detecting malingering and defensive responding on the MMPI-2 in a forensic inpatient sample. J Pers Assess 1994;78:69–86.
18. Cornell DG, Hawk GL. Clinical presentation of malingerers diagnosed by experienced forensic psychologists. Law Hum Behav 1989;13:375–83.
19. Heinze MC. Developing sensitivity to distortion: utility of psychological tests in differentiating malingering and psychopathology in criminal defendants. J Forens Psychiatry Psychol 2003;14:151–77.
20. Fazel S, Danesh J. Serious mental disorder in 23,000 prisoners: a systematic review of 62 surveys. Lancet 2002;359:545–50.
21. Kucharski LT, Toomey JP, Katarzna F, Duncan S. Detection of malingering of psychiatric disorder with the Personality Assessment Inventory: an investigation of criminal defendants. J Pers Assess 2007;88: 25–33.
22. Walters GD. Coping with malingering and exaggeration of psychiatric symptomatology in offender populations. Am J Forensic Psychol 2006; 24:21–40.
23. Walters GD, Rogers R, Berry DTR, Miller HA, Duncan SA, McCusker PJ, et al. Malingering as a categorical or dimensional construct: the latent structure of feigned psychopathology as measured by the SIRS and MMPI-2. Psychol Assess 2008;20:238–47.

Additional information and reprint requests:
Glenn D. Walters, Ph.D.
Psychology Services
FCI-Schuylkill
PO Box 700
Minersville, PA 17954
E-mail: gwalters@bop.gov