

**Marshall University**

---

**From the Selected Works of Marc A. Lindberg Ph.D.**

---

Winter January 1, 2014

# The Many Faces of Malingering and Participant Response Strategies: New Methodologies in the Attachment and Clinical Issues Questionnaire (ACIQ)

Marc A. Lindberg, Ph.D.



This work is licensed under a [Creative Commons CC BY-SA International License](https://creativecommons.org/licenses/by-sa/4.0/).



Available at: [https://works.bepress.com/marc\\_lindberg/11/](https://works.bepress.com/marc_lindberg/11/)

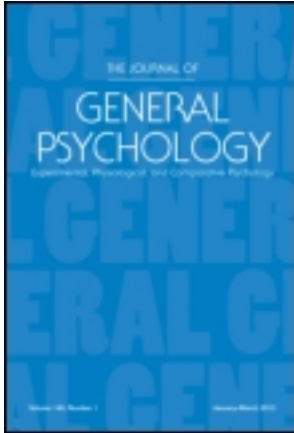
This article was downloaded by: [Marc A. Lindberg]

On: 09 April 2014, At: 14:18

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954

Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## The Journal of General Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/vgen20>

### The Many Faces of Malingering and Participant Response Strategies: New Methodologies in the Attachment and Clinical Issues Questionnaire (ACIQ)

April Fugett<sup>a</sup>, Stuart W. Thomas<sup>a</sup> & Marc A. Lindberg<sup>a</sup>

<sup>a</sup> Marshall University

Published online: 01 Apr 2014.

To cite this article: April Fugett, Stuart W. Thomas & Marc A. Lindberg (2014) The Many Faces of Malingering and Participant Response Strategies: New Methodologies in the Attachment and Clinical Issues Questionnaire (ACIQ), *The Journal of General Psychology*, 141:2, 80-97, DOI: [10.1080/00221309.2013.866538](https://doi.org/10.1080/00221309.2013.866538)

To link to this article: <http://dx.doi.org/10.1080/00221309.2013.866538>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages,

and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# The Many Faces of Malingering and Participant Response Strategies: New Methodologies in the Attachment and Clinical Issues Questionnaire (ACIQ)

APRIL FUGETT  
STUART W. THOMAS  
MARC A. LINDBERG  
*Marshall University*

---

**ABSTRACT.** Four studies created malingering and response bias scales for a new test battery, the Attachment and Clinical Issues Questionnaire (ACIQ). In the first calibration study, a new approach to identifying fake good and fake bad respondents was outlined. In Study 2, this scale was cross validated in a within-subjects design that also found only weak correlations between the scales of the ACIQ and measures of social desirability. The third study developed a method violator scale (one who responds randomly to the content of the scales due to carelessness, low IQ, etc.). It was tested by Monte Carlo and empirical studies. The fourth study combined the two cross validation studies to offer clear cutoffs for the practitioner. These studies successfully led to the creation of malingering and response bias scales for the ACIQ and also introduced new methods that could be adapted to other instruments.

**Keywords:** ACIQ, attachment measurement, malingering, social desirability, response bias

---

TOO OFTEN, PSYCHOLOGISTS employ test and assessment instruments that do not deal with issues of malingering and response bias; issues that can potentially invalidate instruments and/or diagnosis. This seems particularly true in the area of adult attachment (Cassidy & Shaver, 2008). When one wishes to use a self-report instrument in forensic, clinical, or other applied settings, one must first deal with the many different facets of response bias. Response bias has been defined by *The Standards for Educational and Psychological Testing* (1999, p. 181) as “A test taker’s tendency to respond in a particular way or style to items on a test . . . that yields systematic, construct-irrelevant error in test scores.” From this broad definition, it can be seen that problems from response biases could be produced by

---

*Address correspondence to Marc A. Lindberg, Department of Psychology, Marshall University, Huntington, WV 25755, USA; Lindberg@Marshall.edu (e-mail).*

a variety of sources. Thus, if a self-report questionnaire is to achieve scientific and clinical integrity, it must contain scales that can deal with these potential sources of error. It was the purpose of the present investigation to test new scales to be used in the Attachment and Clinical Issues Questionnaire (ACIQ) (Lindberg, Fugett, & Thomas, 2012; Lindberg & Thomas, 2011) as well as to provide potential methodologies for determining response bias for other psychological tests.

The ACIQ was founded upon (1) observations of clinical patients in recovery groups in hospitals, out-patient recovery groups, and 12 step groups, (2) the clinical and self-help literature dealing with addictions, depression, and personality disorders, and (3) the attachment theories proposed by Bowlby (1958; 1960; 1969/1982, 1973), Ainsworth (1964, 1989), and Ainsworth, Blehar, Waters, and Wall (1978). Based upon these foundations and preliminary psychometric studies, 29 scales were crafted. These scales and representative items can be seen in Lindberg & Thomas (2011). These scales were designed to measure secure, avoidant, ambivalent, and codependent/enmeshed attachments to mother, father, and partner. In addition, the ACIQ also has the following scales: Abuser, Anger, Anxiety, Control, Denial of feelings, Family rigidity versus chaos, Family suppression of feelings, Jealousy, Rumination, Peer relations, Perfectionism, Religious practices, Sexual arousal, Sexual intimacy with partner, Shame, Mistrust, and Withdrawal versus engagement. Initial validity studies (Lindberg & Thomas, 2011) have shown that the scales load on seven factors that represent attachment figure rather than attachment pattern or styles, and that the scales have good test-retest reliability. Further, Lindberg, Fugett, and Thomas (2012) have shown that the scales predict the gold standard of attachment theory, to whom one turns in times of stress, as well as showed psychometric superiority to the Experiences in Close Relationships scales of Brennan, Clark, and Shaver (1998) in measures of partner satisfaction, father warmth, and mother warmth.

One of the most troubling sources of response bias to tests like this and to clinicians has been termed "Malingering." Malingering has been defined in the DSM-IV (American Psychiatric Association, 1994), as the intentional production of false or grossly exaggerated symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution or obtaining drugs. Although this type of intentional response bias is important, Faust and Ackley (1998) have reasoned that response bias is more complex than this simple DSM-IV classification and can be seen as coming from several sources (*Standards for Educational and Psychological Testing*, 1999).

To illustrate the kinds of participant response biases that could invalidate ACIQ profiles, let us consider a hypothetical family of four who were court ordered into therapy. Suppose that the husband is a "super salesman" whose general nature is to categorize things, including himself, in ways that are socially desirable. He does not intentionally try to fake "good;" it just comes as second nature to him. His wife, on the other hand, has a strong tendency to try and convince others

that she is better than she really is, and although she suffers from a variety of psychopathologic maladies, she intentionally fakes “good” to the clinician and tries to act as a model wife and mother devoid of any psychological issues at all. Fitting the classic DSM notion of malingering, Adolescent 1 intentionally presents a bad, or psychopathologic front, as a cry for help. The second child, Adolescent 2, has become tuned out, and avoids discussions with everyone connected to the family and therapist and is very uncooperative. Would these four individuals all fill out the ACIQ in a valid fashion? Without proper response bias scales and measures, we would have no way of knowing.

From the example, the father represents one potential type of unintentional response bias that could weaken the validity of a test score. This type of bias has been termed “social desirability.” “Social desirability” is defined as an attempt to answer questions in ways that conform to typical social expectations that would be popular or acceptable with one’s group or target audience. Crowne and Marlowe (1964) developed the classic scale that was empirically demonstrated to be sensitive to this kind of response bias. (See McCrae & Costa (1983) and Smith & Ellingson (2002) for additional discussions.) One purpose of the present investigation was to test whether social desirability negatively affected the validity of the ACIQ or its subscales by testing if the Marlowe-Crowne test (Crowne & Marlowe, 1964) accounted for a significant amount of variance on the scales of the ACIQ. Further, could this test of “unconscious” or unintentional social desirability account for a majority of variance in malingering scales developed here that were designed to detect conscious intentional response bias (malingering as defined by the DSM-IV, and also referred to as faking)?

Malingering as a source of response inaccuracy has been a source of concern for most clinicians. These types of response biases are seen in the examples of the mother and Adolescent 1. In contrast to the mother who was attempting to present a “good front,” or to use the present terminology “fake good,” the opposite would hold for Adolescent 1 who was presenting negative symptoms or “faking bad.” A traditional way in which these types of intentional malingering have been tested is through the use of a fake good/bad design where participants are told to engage in response bias strategies. Although it is necessary to develop internal measures that can detect this type of potential deception, such an instrument must also be sensitive to the fact that it might be clinically dangerous to label a true responder with significant psychological issues a “faker” and not engage in appropriate treatment. Thus, although response bias scales must be used in clinical and applied situations, they must have very low false positive rates if one is to preserve safety, clinical sensitivity, and trust.

Another variable that can potentially confound validity is demonstrated by Adolescent 2, who did not take the test seriously. These types of validity violations will be referred to in this article as “Method Violations.” Thus, if one filled out the test carelessly or randomly, then one would expect these participants to score high on a Method Violation validity scale.

The first study attempted to develop scales that would test for the intentional malingering of response bias and still be conservative in falsely accusing an accurate and truthful responder that they were “faking.” The second study was a within-subjects design and attempted to provide a cross validation of the response bias scale and also test whether social desirability would impact response validity negatively. This study also attempted to get at the question of whether response bias methodologies versus social desirability tests measure different constructs in the context of the ACIQ. The third study presented a second cross validation of the response bias scales and also attempted to create and test a Method Violators scale both in terms of Monte Carlo designs and empirically by instructing participants to respond by just filling in the blanks quickly without regard to the particular questions. The fourth study combined the two cross validation studies to offer clear cutoffs for the practitioner.

### Study 1

The purpose of Study 1 was to develop scales from the ACIQ that could detect those who were intentionally using a negative response bias or those who were using a positive response bias. It can be recalled that participants are instructed to rate the questions on a four point scale going from 1 = never, 2 = sometimes, 3 = often, and 4 = always.

The first study will be referred to as the calibration study because its purpose was to create a response bias scale from items on the ACIQ. Instructing examinees to fake good (FG) or fake bad (FB) or answer accurately and truthfully (AT) provided an empirical means of identifying test items most susceptible to the above types of response bias. In the case of Likert-type items, it was reasoned that one should be able to identify monotonic target items where the FB examinees produce significantly lower mean responses than AT examinees who in turn respond significantly lower than FG examinees. These will be referred to as ascending items (ASC) because of the following monotonic order of significant differences in means on the items for the groups  $FB < AT < FG$ . Descending (DSC) monotonic target items, on the other hand, would be those items on the ACIQ where the following groups would have the following significantly different orderings of means  $FB > AT > FG$ . (A good mnemonic to help one remember the abbreviations is to say that one ascends to good (ASC) and one descends from bad (DSC)).

The total ASC score would be the sum of all ASC item calculations added together. DSC total scores would be similarly calculated. Once identified, it is hypothesized examinees using a negative response bias (faking bad) would have a low total score for ASC items and a high total score on DSC items. Those using a positive response bias (faking good), on the other hand, would have high total scores for ASC items and low total scores on DSC items. Answering truthfully should typically lead to intermediate total scores on both ASC

and DSC items. If one reversed the scoring for DSC items, one could then combine the ASC and reversed DSC items into one scale where low scores indicate a negative response bias and high scores would indicate a positive response bias providing the clinician with something easily spotted on individual profiles.

Because of the needed conservatism in misdiagnosing response bias, it was determined to use the controls as the basis for establishing the final cutoffs on the response bias scale. We determined that anything more than a 5% false positive error rate was unacceptable; and therefore, arranged the cutoffs such that 95% in the AT control group would be correctly placed in the control group on the response bias scale.

## Method

### Participants

The Calibration study had 277 participants with 23 males and 64 females in the Fake Good condition, 25 males and 60 females in the Fake Bad group, and 44 males and 57 females in the control group. Four did not indicate gender. These were introductory psychology students who were earning extra credit for their introductory psychology course. This and all studies in this article were approved by the Institutional Board of Review.

### Procedures

The participants were instructed to first read the instructions on the ACIQ. Those in the control conditions then filled out the ACIQ. Those in the FB condition were then read a statement asking them to give a slightly worse impression of themselves and those in the FG condition were asked to give a slightly better impression of themselves. These instructions can be seen in the Appendix. After all participants were finished, they were debriefed with a more complete explanation of the procedures, and a brief lecture on psychometrics, test construction, and attachment theory.

### Calibration Technique

The steps in the calibration study can be seen in Figure 1 of this article. Mean scores on all 236 ACIQ items were calculated for the positive response bias (faking good or FG) sample, the negative response bias (faking bad or FB) sample, and the control (AT) sample. In step one, items were selected if their mean scores had either of the following patterns:  $FB < AT < FG$  (ASC) or  $FB > AT > FG$  (DSC). In step 2, those monotonic items where the difference in FB/FG means failed to yield a two-tailed  $t$ -test  $p < .05$  were then deleted. Response bias scores were calculated



1. First get monotonic items from the ACIQ where the three groups differed in the following orders:

ASC	DSC
FB<AT<FG	FB>AT>FG

2. Only those items where there was a significant difference ( $p < .05$ ) between the scores for FB scale and FG scale in the predicted order were then used.

3. These ASC item scores and the reversed DSC item scores are summed to obtain response bias scores.

4.*	Response bias scores	Response bias scores
	Lowest 2.5%	Highest 2.5%

**FIGURE 1. A description of the steps involved in the calibration study. \*If one is in the lowest 2.5% of the response bias scale, then one is predicted to have faked bad and if one is in the highest 2.5%, then one is predicted to have faked good. The false positives for response bias should be no greater than 5% of the AT control group. It is also important to note that these items and scales are embedded within the ACIQ itself. These scales have been normed across the 1221 participants in Lindberg and Thomas (2011). The scores come out as a standard score on the profile's printout for each participant along with their standard scores for each of the 29 scales.**

by adding all ASC items and the reverse scored DSC items together. To meet the conservative criterion of not incorrectly classifying the control participants, the cutoff scores for classification as faking were set at the 2.5-th and 97.5-th control AT sample percentiles of the Response bias scale sum. Thus, no more than 5% of the AT participants would be false positives for response bias given the above assumptions.

It is again useful to recall that the classification process was AT biased in that all participants were initially classified as belonging to the AT group. Only those participants whose response bias scores were in the extreme lowest or highest 2.5 percent were classified as faking. Thus, no more than 5 percent of the AT participants would be false positives for response bias.

## Results

The calibration sample yielded 51 ASC items and 113 DSC items with  $p < .05$ . The DSC items were reversed and combined with the ASC items to yield a 164 item response bias scale.

## Study 2: Cross Validation

The first purpose of the second study was a methodological replication. We decided to use a within subjects design wherein all participants would take the ACIQ under standard conditions, and then take it again under either fake bad or fake good instructions. This within subjects study was designed to yield data addressing the difference between an examinee's control (AT) response and his/her faked response for each item. It was also designed to see how effective the scales were in diagnosing when participants were using a response bias versus when they were taking it under AT instructions.

The second purpose of the present study was to examine the problem of social desirability as defined by Crowne and Marlowe (1964) and to compare and contrast social desirability with faking. As Smith and Ellingson (2002) have so eloquently noted, the debate surrounding the construct of social desirability has continued for more than 30 years, with some treating social desirability as a substantive trait (Block, 1965; McCrae & Costa, 1983; Nicholson & Hogan, 1990) and others treating it as response distortion that alters the ability to interpret personality trait information (Edwards, 1957; Griffith, 1997). Placed in this context, we must ask "Does one's tendency to appear socially desirable negatively impact the content validity of the 29 scales of the ACIQ?" Going back to the super salesman illustrated in the general introduction, we have to ask if the validity of his profile would be adversely affected by his automatic tendency to place himself and most everything else in a positive light. Because it has also been a concern as to whether response bias methodologies tap social desirability constructs, the present study also compared the response bias scale developed here with the Marlowe-Crowne scale. To place this in the context of the example outlined in the introduction, would the super salesman father's positive spin approach be qualitatively different than the mother who was deliberately trying to appear better than she really was? Thus, we also attempted to shed light on the questions posed by Smith and Ellingson (2002) regarding possible differences between social desirability versus response bias.

## Method

### Participants

The Fake Good group had 11 males and 21 females and the Fake Bad group had 13 males and 15 females. The participants were college students participating for extra credit in introductory psychology.

### Procedure

When the participants first took the ACIQ, they were given the standard instructions. They then took the Marlowe-Crowne (1964) test of social desirability.

A week later, half of the participants were instructed to take it again under fake good instructions and the other half took it again under fake bad instructions. When they took the ACIQ under the response bias instructions, they were given the same instructions as in Study 1, but were first told “In creating a test like the one you previously took, it is important to determine when people are trying to look better than they really are and when they try and look worse. The purpose of your taking this test again is to see which questions are most sensitive to people’s needs to look better than they really are and to look worse than they really are.” They were then given either the fake bad or fake good instructions.

## Results and Discussion

### Response Bias Data

The first set of analyses examined the cross validation data for the Response bias scale. The first analysis was performed on the percentage of calibration items which showed significant changes ( $p < .05$ ) under the response bias conditions in cross-validation. Table 1 shows the percent of items that significantly differed under FB versus AT instructions, under FG versus AT instructions, and under FB versus FG instructions.

The next analyses applied the Response bias score determined from the calibration sample to the present cross-validation sample. This was to test how accurately the response bias scale from the calibration study could predict when one was taking the test accurately and truthfully versus when they were faking. These data can be seen in Table 2.

From Table 2, it is apparent that the scale exceeded the psychometric criteria we had established under this cross validation procedure. Both the item calculations and the classifications based upon those classifications easily distinguished when one was employing response bias versus when he/she was not. It can be seen that it was apparently easier to detect a negative response bias, as 93% of the

**TABLE 1. Percent of Items Showing Significant Differences Between the Fake Bad (FB) and Accurate and Truthful (AT) Controls, the Fake Good and AT Controls, and the FB and FG Groups**

Item type	ASC ( $N = 51$ )	DSC ( $N = 113$ )
Difference		
FB/AT	96%	87%
FG/AT	92%	79%
FB/FG	100%	96%

**TABLE 2. Percent of Participants from the Fake Bad, Control, and Fake Good Groups Correctly Identified Based Upon the Response Bias Scale**

From This Group	Classified As			Sample <i>N</i>
	Fake-bad	Control	Fake-good	
Fake-bad	92.8% 26	7.2% 2	0.0% 0	28
Control	1.7% 1	95.0% 57	3.3% 2	60
Fake-good	0.0% 0	21.9% 7	78.1% 25	32
Total	27	66	27	120

FB participants were correctly identified as faking bad whereas 78% of the FG participants were correctly identified as using a positive response bias. Given the cutoff procedures, the expected 95% of the AT control participants were correctly identified.

### Social Desirability Data

The second purpose of Study 2 was to shed light on whether measures of social desirability tap the same constructs as fake good methodologies and if so, whether social desirability correlates strongly with the scales of the ACIQ. The answer to the question on whether the two scales measure the same thing was analyzed through calculating the correlation between the Marlowe-Crowne test and the ACIQ response bias scale. This correlation was ( $N = 58$ )  $r = .33, p < .01$ . However, because the amount of variance that the Marlowe-Crowne accounted for in the response bias scale was only 11%, it is obvious that the two scales did not measure the same construct, agreeing with the more substantive studies of McCrae and Costa (1983) and Smith and Ellingson (2002). Here, we directly observed that faking using the above measures and social desirability were not the same things and should not be confused, thus providing evidence of discriminant validity between the two constructs.

In an attempt to explore whether social desirability affected the 29 scales of the ACIQ, correlations were performed between the 29 scales and the Marlowe-Crowne. From Table 3, it can be seen that the correlations averaged ( $n = 58$ ),  $r = .22$ , meaning that this test of social desirability typically accounted for less than 5% of the scale variances. These results provide more evidence of discriminant validity between social desirability and faking. In summary, the ACIQ was fairly

**TABLE 3. ACIQ Scale Correlations With the Marlowe-Crowne**

ACIQ Scale	Marlowe-Crowne
1. Abuser	-.33
2. Ambivalent Father	-.15
3. Ambivalent Mother	-.32
4. Ambivalent Partner	-.09
5. Anger	-.14
6. Anxiety	-.24
7. Avoidant Father	-.24
8. Avoidant Mother	-.36
9. Avoidant Partner	-.13
10. Cod/enmeshed Mother	-.06
11. Cod/enmeshed Father	-.27
12. Cod/enmeshed Partner	-.43
13. Control	-.28
14. Denial	-.31
15. Family Rigidity v Chaos	.10
16. Family supp. of feelings	-.24
17. Jealousy	-.32
18. Obsessive-Compulsive	-.29
19. Peer Relations	-.29
20. Perfectionism	-.10
21. Religion	.11
22. Sexual Arousal	-.36
23. Secure Father	.04
24. Secure Mother	.21
25. Secure Partner	-.12
26. Shame	-.24
27. Sexual Intimacy	-.30
28. Mistrust	-.28
29. Withdrawal v Engagement	-.09

independent of social desirability effects, and is therefore fairly immune from unconscious forms of social desirability as defined and examined by Crowne and Marlowe (1964).

### **Study 3: Identifying Method Violators and Further Cross Validations of the Response Bias Scales**

The next problem we addressed was the problem of the Method Violator. A method violator's responses do not relate to the content of test items because the examinee is sleepy, inattentive, uncaring, misunderstands, or is uncooperative. In the case of the ACIQ, how might one recognize method violators? The purpose of

the present study was to develop a Method Violator scale as well as to provide an additional cross validation study for the Response bias scale.

If examinees with secure attachments to their mothers attend to the items, their responses to the secure mom scale should differ from their responses to the avoidant mom scale. They should have greater propensities for the higher-valued responses and lesser propensities for lower-valued responses for items in the secure mom scale; the propensities should be reversed for the avoidant mom scale. On the other hand, MVs securely attached to their mothers would be inattentive to item content, so there is no a-priori reason to expect their response propensities to be affected by the differing contents of the two scales.

Based upon these considerations, we conceptualized the method violator (MV) as one whose response to every ACIQ item would be an independent random selection from a fixed response distribution (FDR) where  $p_1$ ,  $p_2$ ,  $p_3$ , and  $p_4$  are the fixed probabilities of their Likert scale responses 1, 2, 3, and 4. Although the MV's responses would differ from item to item, the expected value of the response to each item would be the same, and consequently all means for the 29 ACIQ scales would have the same expected value. For example, if a normal responder scored high on the Secure Mom scale, they would be expected to score low on the Avoidant Mom scale. Similarly, if they scored high on Peer relations, they would be expected to score low on the Withdrawal scale. Scoring high on the Abusive scale would most often entail higher scores on the Anger scale. Thus, variation and patterns would be expected between the different scale means for true responders. The variability of the scale means for the MV, on the other hand, should be expected to be smaller because such scale mean differences would be due entirely to random differences on the responses across the 29 scales.

The variability of an examinee's responses to the 236 ACIQ items can be measured by the total sum of squared deviations (TSS) of the responses around their mean. Analogous to a 1-factor ANOVA, the TSS is equal to the combined within-scale sums of squares (WSS) plus the sum of the squares of the scale means (BSS). If the preceding model of the MV is appropriate, most of a MV's response variability will be within-scale variability; hence, for any given TSS we would expect MVs to have larger values of WSS (since  $WSS = TSS - BSS$ ) than examinees who attend to item content. We propose the ratio  $L = WSS/TSS$  as a means of identifying MVs; large values of  $L$  would be consistent with the low variability of scale means expected of MVs.

The variance of an examinee's fixed response probabilities (FRD) will determine the observed values of the TSS and WSS. The expected value of the TSS is 236 (the number of items in the ACIQ) times the variance of the FRD, and the expected value of the WSS is 207 times the variance of the FRD. An approximate expected value of  $L$  for a MV is therefore 0.88, implying that around 88% of the overall variability of a MV's responses is within the scales, and around 12% is due to variability of the 29 scale means. (See Searle (1971), Theorem 1).

In summary, the purpose of the present study was twofold. First, it was designed to test empirically the above statistical model of the method violator and develop a scale for detecting method violators based on it. Second, it was designed to provide another cross validation of the fake bad and fake good scales. This was done by instructing four different participant groups to either fake bad, fake good, respond randomly, or take the ACIQ under the standard conditions. Could our scales detect who belonged to what group?

## Method

### Participants

A total of 110 introductory students participated in this study to earn extra credit. The following were the breakdowns of group by numbers from each sex: There were 6 males, 9 females, and 5 not marking this item in the FB condition. There were 2 males and 21 females in the Control group with 10 not indicating sex on their forms. There was one male and 30 females in the FG condition with 2 not indicating sex. There were 24 participants in the random group, and coding for sex was not possible for this group given their instructional set.

### Procedures

The participants signed up for group testing sessions, and the sessions were randomly assigned to one of four conditions. In the Control condition, participants were instructed to fill out the ACIQ under the standard instructions. The FB group was given the previously mentioned fake bad instructions except this time they were told to look much worse and “psychologically unhealthy” rather than just give a “slightly worse impression of you.” The third group, the FG group, was given the previously mentioned fake good instructions except this time they were told to “look much better and totally psychologically healthy rather than to give a slightly better impression of you.” The fourth group, the random group (RNDM), was told to answer the questions randomly. If they wanted, they could even just fill in as many blanks as there were questions on the test. (Here it should be noted that the experimenter observed that the participants in this condition seemed to not read many of the questions and just marked a varied array of responses on the answer sheets.)

### Results for the Random Group, and Comparisons to Monte Carlo Studies

These examinees' mean *L* was 0.85 with a standard deviation of .088. The *L* values ranged from .77 to .91 plus an outlying value of .47. If the outlier is omitted, the mean was .87 and the SD ( $n = 23$ ) = .035. This mean value was very close to the expected value of .88.

**TABLE 4. Observed and Simulated L Values for Various Percentiles of Fixed Response Probability Distributions (FRPD)**

FRPD	Observed ( <i>n</i> = 23)	Simulated ( <i>n</i> = 1,000)
99th pctile	.92	.93
95th	.91	.91
90th	.91	.90
75th	.89	.88
50th	.87	.86
25th	.85	.84
10th	.83	.82
5th	.81	.81
1st	.77	.78
Mean	.87	.86
SD	.035	.030

Data from the 23 random group examinees (the outlier was omitted) were combined to obtain a single overall FRD for the random responders for generating 1,000 sets of 236 simulated responses. The 1,000 sets of 236 responses were then scored according to the ACIQ scales and the value of L was determined for each set. The mean L and standard deviation of the 1000 L values was determined. The observed and simulated L distribution data can be seen in Table 4. It is clear that the observed and simulated values of L correspond closely. Data generated from the model offered a good description of the observed distribution of L values with no data characteristic differing by more than .01 between the observed and simulated distributions.

Here it must be emphasized that the MV scale’s ability to detect a variety of different types of invalid response patterns could be regarded as a liability. Although Faust and Ackley (1998) have made an excellent case that invalid profiles could come from the miss administration of the test, or the participant’s carelessness in filling it out, or from a low IQ, the present scale can’t determine why the participant did not respond to the content of the scales in a valid fashion. However, the purpose of the present scale was not to aid the forensic investigator in discerning what type of method violations occurred. Rather, its purpose was to offer a scale that was useful in helping guide the clinician and researcher’s confidence in particular profiles as offered by the ACIQ.

**Results for Correctly Assigning to Group by the Scales**

When the 95% rule for control placement was applied, there was again a good fit of scale score to group. These data can be seen in Table 5. Table 6 shows the data



**TABLE 5. Percent of Participants in the Fake Bad (FB) Group, the Control Group, the Fake Good (FG) Group, and the Method Malingering Group Correctly Assigned to Condition Given Their Scores on the Basis of Scale Scores**

	FB	Control	FG	Method Malingering
FB	13 65%	3 15%	0 0%	4 20%
Control	0 0%	31 93.94%	1 3.03%	1 3.03%
FG	0 0%	8 34.78%	14 60.87%	1 4.35%
Random	0 0%	1 4.17%	5 20.83%	18 75%

after the Method Violators were taken out. Thus, in line with the control biased approach, 97% were correctly classified into the control group, 64% correctly into the FG group, 81% into the FB group.

#### Study 4: The Studies Combined and All Scales Tested

The final set of analyses combined the present study with the within-participants validation study in an effort to provide a more general test of each of the scales that were developed in this series. A second purpose of this analysis was to provide a standard cutoff that would be best across all the studies and could be used on an individual basis for participants taking the test. These data can be seen in Table 7. After one removes the method violators, the results can be seen

**TABLE 6. Percent of Participants Correctly Assigned to the Fake Bad (FB) Group, the Control Group, and the Fake Good (FG) Group After the Method Malingers Were First Taken Out**

	FB	Control	FG
FB	13 81.25%	3 15%	0 0%
Control	0 0%	31 96.88%	1 3.03%
FG	0 0%	8 36.36%	14 63.64%

**TABLE 7. The Cross Validation Studies Combined Under a Standard Cutoff**

	FB	Control	FG	Random
FB	40 83.33%	8 16.67%	0 0%	4 20%
Control	4 4.3%	87 93.55%	2 2.15%	1 3.03%
FG	0 0%	16 29.09%	39 70.91%	1 4.35%

in Table 8. It was found that the Fake bad scale correctly identified 89% of those using a negative response bias and 74% of those using a positive response bias while leaving 94% of the controls in the control group. It should be cautioned, however, that we did not sample anyone suffering from a psychopathology. Thus, one should be very cautious in generalizing these results to such populations. How, for example, would one distinguish between one suffering from psychopathology and a “fake bad” responder? This caveat is extremely important for clinicians to consider and should be dealt with in further studies. It should be pointed out, however, that screening prisoners with the “fake good” scale has been found to improve the amount of variance that the scales of the ACIQ accounted for with criminal populations (Lindberg, Lounder, & Fugett, in press).

**TABLE 8. After the Method Malingering Responders Were Removed Through the Use of the Method Malingering Scale, the Results for the Cross Validation Studies Combined Were as Follows**

	FB	Control	FG
FB	39 88.64%	5 11.36%	0 0%
Control	4 4.35%	86 93.48%	2 2.17%
FG	0 0%	14 26.42%	39 73.58%

## Summary and Conclusions

The purpose of these studies was to deal with the problem of any self-report measure; that examinees can often present self-images either better or worse than how they honestly perceive themselves. They might do this either unconsciously as measured by the Marlowe-Crowne test, or they might do it consciously as in response bias designs in an attempt to dupe the clinician, researcher, or forensic psychologist. They might also not take the test seriously, answering the questions carelessly or randomly. Alternatively, the test might not have been properly administered or the examinee might not have had the intellectual ability to fill it out properly. Lacking assessment of these possibilities makes one's confidence in using such a measure in a clinical or research setting a matter of concern. It should be pointed out that this was an exploratory investigation of the scales and instrument in terms of response bias, and it should be emphasized that there is a need for future studies with this instrument that include a clinical comparison group.

The present studies presented some new logic for scale development. In the first three studies, the developed scales were able to detect 89% of those "faking bad," 74% of those "faking good," and best of all, retain 94% of the participants in the control group. The Method Violation scale was able to detect 75% of those instructed to respond without regard to item content, and Monte Carlo studies matched empirical results. Thus, these scales were successful in application. Further, they provided conservative estimates of malingering in control populations. As was outlined in the introduction, this is of importance in clinical, forensic, and other applied settings where it can be damaging to term a true responder a "fake." Such false positives could have the result of leaving a person who is hurting from some significant issue without treatment. Without such scales and tests, the attachment and clinical scales of the ACIQ, or any such instrument, would have limited usefulness in research or practice.

## AUTHOR NOTES

**April Fugett** is an Associate Professor at Marshall University. Her research interests include technology, language, reading comprehension, memory, statistics, methodology, and the influence of popular culture in terms of memory and language. **Stuart W. Thomas** was a professor in the Department of Psychology at Marshall University. He received his Ph.D. from Michigan State University specializing in quantitative psychology. Dr. Thomas died of ALS on May 26, 2003. He served Marshall University throughout his career. His brilliance was accompanied by a warm heart and loving compassion for students and colleagues alike. **Marc A. Lindberg** is a Professor at Marshall University. His current specialty is in developmental psychopathology, and is specializing in dynamic systems interventions based on the ACIQ diagnostic to deal with addictions and recidivism in prisoners.

## ACKNOWLEDGMENTS

The authors would like to thank Cecil Reynolds for his criticisms of earlier drafts of this article. Stuart W. Thomas died of ALS in the Spring of 2003.

## REFERENCES

- Ainsworth, M. D. S. (1964). Patterns of attachment behavior shown by the infant in interaction with his mother. *Merrill-Palmer Quarterly*, 10, 51–58.
- Ainsworth, (1989). Attachments beyond infancy. *American Psychologist*, 44, 709–716.
- Ainsworth, M. D. S., Blehar, M., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation*. Hillsdale, NJ: Erlbaum
- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Press.
- Block, J. (1965). *The challenge of response sets*. New York: Appleton-Century-Crofts.
- Bowlby, J. (1958). The nature of the child's tie to his mother. *International Journal of Psycho-Analysis*, 39, 350–373.
- Bowlby, J. (1960). Grief and mourning in infancy. *Psychoanalytic Study of the Child*, 15, 3–39.
- Bowlby, J. (1969/1982). *Attachment and loss: Vol. 1. Attachment*. New York, NY: Basic Books.
- Bowlby, J. (1973). *Attachment and loss: Vol. 2. Separation*. New York, NY: Basic Books.
- Bowlby, J. (1973). *Attachment and loss: Vol. 3. Loss*. New York, NY: Basic Books.
- Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult attachment: An integrative overview. In J. A. Simpson & W. S. Rhodes (Eds.), *Attachment theory and close relationships* (pp. 46–76). New York, NY: Guilford Press.
- Cassidy, J., & Shaver, P. R. (2008). *Handbook of attachment: Theory, research, and clinical applications* (2nd ed.) New York: The Guilford Press.
- Crowne, D. P., & Marlowe, D. (1964). *The approval motive: Studies in evaluative dependence*. New York, NY: John Wiley & Sons.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: The Dryden Press.
- Faust, D., & Ackley, M. A. (1998). Did you think it was going to be easy? Some methodological suggestions for the investigation and development of malingering detection techniques. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation* (pp. 1–54). New York, NY: Plenum Press.
- Griffith, R. (1997). *Faking of non-cognitive selection devices: Red herring is hard to swallow*. Unpublished doctoral dissertation, University of Akron, Akron, OH.
- Lindberg, M. A., Fugett, A., & Thomas, S. W. (2012). Comparing Measures of Attachment: "To whom one Turns in Times of Stress," Parental Warmth, and Partner Satisfaction. *Journal of Genetic Psychology*, 172 (4), 329–352.
- Lindberg, M. A., & Thomas, S. (2011). The Attachment and Clinical Issues Questionnaire (ACIQ): Scale Development. *Journal of Genetic Psychology*, 173 (1), 41–62.
- McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, 51, 882–888.
- Nicholson, R. A., & Hogan, R. (1990). The construct validity of social desirability. *American Psychologist*, 45(2), 290–292. doi: 10.1037/0003-066X.45.2.290
- Lindberg, M. A., Fugett, A., & Lounder, L. (in press). The attachment and clinical issues questionnaire: A new methodology for science and practice in criminology and forensics. *International Journal of Offender Therapy and Comparative Criminology*.

- Searle, S. R. (1971). *Linear models*. New York, NY: John Wiley & Sons.
- Smith, D. B., & Ellingson, J. E. (2002). Substance versus style: A new look at social desirability in motivating contexts. *Journal of Applied Psychology*, 87, 211–219.
- Standards for Educational and Psychological testing*. (1999). Washington, DC: American Educational Research Association.

*Original manuscript received June 28, 2013*

*Final version accepted November 13, 2013*

## **Appendix**

### **Faking Instructions**

#### **Fake Bad**

You are in the condition designed to study cases when people try and look worse than they really are. When you take this test, suppose that you are involved in an insurance claim. Suppose that a Psychologist gives you this test, and your answers will determine whether or not you will be able to qualify for your rightly earned settlement with the insurance company if you look as bad as possible. In other words, to get your justly owed insurance claim, it is very important for you to look a little bit worse than you really are. When you take this test, try and place your worst foot forward, and respond to these questions as though looking bad to the psychologist and others is the most important thing that you do. In summary, when you answer these questions this time, answer them from a slightly worse impression of you.

#### **Fake Good**

You are in the condition designed to study cases when people try and look better than they really are. When you take this test, suppose that you are involved in an insurance claim. Suppose that a Psychologist gives you this test, and your answers will determine whether or not you will be able to qualify for your rightly earned settlement with the insurance company if you look as good as possible. In other words, to get your justly owed insurance claim, it is very important for you to look a little bit better than you really are. When you take this test, try and place your best foot forward, and respond to these questions as though looking good to the psychologist and others is the most important thing that you do. In summary, when you answer these questions this time, answer them from a slightly better impression of you.