JOURNAL OF PERSONALITY ASSESSMENT, *88*(3), 328–337

Varying Cautionary Instructions, Monetary Incentives, and Comorbid Diagnostic Training in Malingered Psychopathology Research

Jon D. Elhai and Jimmie J. Butcher
Disaster Mental Health Institute
University of South Dakota

Andrea N. Reeves and Shannon N. Baugher
Department of Psychology
University of South Dakota

Matt J. Gray
Department of Psychology
University of Wyoming

Gerard A. Jacobs
Disaster Mental Health Institute
University of South Dakota

Adrienne E. Fricker-Elhai
Department of Psychiatry
University of South Dakota School of Medicine

Terry C. North
Omaha Veterans Affairs Medical Center
Omaha, Nebraska

Paul Arbisi
Minneapolis Veterans Affairs Medical Center
Minneapolis, Minnesota

In this article, we combine two analogue experiments in which we empirically examined three malingering methodological issues in individuals trained and instructed to simulate posttraumatic stress disorder (PTSD) on the Trauma Symptom Inventory (TSI; Briere, 1995). In Experiment 1, we examined TSI scale effects of the following manipulations using a 2×2 design with 330 college students: (a) inclusion or exclusion of cautionary instructions regarding believability of participants' simulation and (b) different financial incentive levels. In Experiment 2, we examined comorbid psychiatric diagnostic training with 180 college students who were either trained to simulate PTSD and comorbid major depressive disorder or trained to simulate only PTSD. Caution main effects were significant for all but two TSI Clinical Scales, incentive main effects and interactions were only significant for one Clinical scale each, and the comorbidity manipulation did not yield any scale differences. We discuss malingering research design implications regarding the use of cautionary instructions, financial incentive levels, and comorbid training.

The simulation (i.e., analogue) research design is the most widely used in malingering research (Rogers, 1997a). This design involves comparing individuals instructed to feign a disorder ("simulators") with honest responders on a particular outcome. Although weaknesses regarding generalizability have been noted, the design's primary strength is its controlled manipulation of experimental and comparison conditions (Rogers, 1997a). Among studies that have empirically investigated malingering detection, simulation designs have often included several procedures aimed at increasing their generalizability. Two of these procedures include (a) cautionary instructions for participants regarding the need to appear believable in simulation and (b) monetary incentives to stimulate motivation and effort (Rogers, 1997b). However, neither procedure has been sufficiently tested for its utility using experimental designs.

A third methodological procedure, conversely, has not been used when training participants to simulate psychopathology. This procedure, training simulators to feign a comorbid mental disorder in addition to the target diagnosis, could potentially enhance the generalizability of analogue malingering designs because disorders frequently co-occur with other diagnoses.

CAUTIONARY INSTRUCTIONS

Cautionary instructions to appear believable in psychopathology simulation are important because successful feigning requires evading detection by a mental health professional. Moreover, individuals encounter large incentives to avoid being detected as malingerers because courts have increased prison sentences for malingerers (Knoll & Resnick, 1999). In simulation research, "cautionary instructions" are different from "coaching" in that coaching involves giving specific information on strategies to avoid detection on test validity scales, whereas cautionary instructions do not include such specific strategic information.

Only one study has examined whether cautionary instructions affect overreported psychopathology among simulators. Viglione et al. (2001) tested the effect of cautioning individuals to maintain a realistic symptom profile when simulating depression or general psychological disturbance. Viglione et al. found a significant effect for cautionary instructions. Compared to simulators not given cautionary instructions, cautioned simulators scored significantly lower on several Minnesota Multiphasic Personality Inventory–2 (MMPI–2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) fake bad validity scales, which produced "large" effect sizes (Cohen's d ranged from .85-1.03). However, no other investigations have assessed the effects of cautionary instructions on symptom overreporting.

MONETARY INCENTIVES

Financial incentives have been utilized in malingering research to enhance simulators' motivation (Rogers, 1997b). These incentives provide an analogue (albeit much smaller in amount) to that of actual malingerers who may receive large financial remuneration in successfully malingering psychopathology (e.g., in worker's compensation, disability payments).

However, studies have varied widely regarding how large of an incentive to provide the most effective simulators in each investigation. Financial incentives typically have ranged from as little as $2 (Martin, Bolter, Todd, Gouvier, & Niccolls, 1993) to $100 (Bagby, Buis, & Nicholson, 1995; Bagby, Rogers, & Buis, 1994; D. D. Blanchard, McGrath, Pogge, & Khadivi, 2003; Nicholson et al., 1997). Amounts frequently used have ranged from about $20 to $50 prizes, often awarded to the best two to five simulators (Bernard, 1990; Elhai, Gold, Frueh, & Gold, 2000; Elhai, Gold, Sellers, & Dorfman, 2001). Thus, there is no clear consensus on how much of a financial incentive to offer, and most important, there is little understanding as to what amount is necessary to result in an experimental effect. Without that knowledge, researchers often make decisions on the use of financial incentives based simply on the availability of research funds (Rogers, 1997b). Few studies have explicitly examined the effect of financial incentives despite their wide prevalence of use in malingering designs.

Two studies have examined financial incentive effects on simulated memory impairment using the Wechsler Memory Scale-Revised (Weschler, 1987), Complex Figure Test (Rey, 1941), Auditory Verbal Learning Test (Rey, 1964), Rey Memory Test (Bernard, 1990), and the Multi-Digit Memory Test (Rey, 1964; Martin et al., 1993). Across these studies, no significant effects have resulted from offering no monetary incentives versus $2 (Martin et al., 1993) or $50 (Bernard, 1990) for the most successful simulators. However, although these studies have examined feigned memory impairment, fabricated psychiatric disturbance has not been explored.

Only one study investigated financial incentive effects on simulated psychiatric disorders. Rogers and Cruise (1998) examined simulated depression and found significant differences on the Structured Interview of Malingered Symptomatology (SIMS) (Smith & Burger, 1997) based on whether participants were informed of a positive incentive for successful simulation ($50) or a negative incentive for unsuccessful simulation (posting names on a school bulletin board). Specifically, negative incentives resulted only in higher SIMS Mood scale elevations (Cohen's $d = .36$), which suggested greater depression symptom overreporting. Thus, although there is some evidence that providing incentives may affect malingered psychopathology, little is known about the differential effects of varying incentive levels.

COMORBID PSYCHIATRIC DIAGNOSES

Many mental health conditions present with significant psychiatric comorbidity. Yet, in analogue studies, investigators have most often only trained their simulators on one target disorder rather than the disorder's comorbid diagnoses. Furthermore, we do not know whether providing additional training on a target disorder's diagnostic comorbidity affects simulation performances.

Posttraumatic stress disorder (PTSD) serves as a unique disorder to investigate this issue. PTSD's significant psychiatric comorbidity has been well-documented (Keane & Wolfe, 1990; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995) and not fully explained by diagnostic symptom overlap (Franklin & Zimmerman, 2001). For example, in the National Comorbidity Survey's (Kessler et al., 1995) PTSD-diagnosed participants, 59% of men and 44% of women evidenced at least three other psychiatric disorders. PTSD's most highly comorbid psychiatric condition across genders was major depressive disorder (MDD; 48% of men, 49% of women), followed by alcohol abuse/dependence (52% men, 28% women), and simple phobia (31% men, 29% women).

Several studies that have investigated malingered PTSD detection have reported high rates of comorbidity in their PTSD patient comparison samples (Bury & Bagby, 2002; Elhai et al., 2000, 2001, 2004, 2002). Yet no research to date has examined whether training simulators in a target disorder and its comorbid diagnoses will alter simulation performances. Given the high rate of comorbidity common to PTSD, we believe this is an important empirical question to explore. Further, depression (PTSD's most highly comorbid psychiatric condition) is a logical disorder to begin empirical examination of this question.

AIMS OF THIS STUDY

Our aims were to explore the effects of providing cautionary instructions and different financial incentive levels on malingering performances, combined in a single analysis (Experiment 1), and to explore whether providing simulators with comorbid psychiatric diagnostic training (in addition to target disorder training) affects psychological test scores during simulation (Experiment 2). Experiment 1 hypotheses were that cautionary instructions would produce large effects, but different financial incentive levels would produce small or no effects. Interactions between methods have not been tested, and therefore, we proposed no specific interaction hypothesis. Further, given that participants were asked to simulate PTSD specifically, we expected that Trauma Symptom Inventory (TSI; Briere, 1995) scales corresponding to the three PTSD symptom criteria (Intrusive Experiences, Defensive Avoidance, and Anxious Arousal scales) would most likely be impacted. Concerning Experiment 2, we had no empirically based hypotheses regarding the results, as no previous studies have examined this issue. However, we expected that participants simulating both PTSD and MDD would score higher on the TSI's Depression scale, explicitly designed to assess for depressed mood and cognitions (Briere, 1995).

In this study, we investigated these methodological issues with regard to simulated PTSD, a commonly and easily fabricated condition with large financial incentives for feigning (Frueh, Hamner, Cahill, Gold, & Hamlin, 2000; Guriel & Fremouw, 2003; McNally, 2003; Resnick, 1997). Further, we explored simulation using the TSI (Briere, 1995), a well-validated (Briere, 1995; McDevitt-Murphy, Weathers, & Adkins, 2005), widely used clinical posttraumatic stress symptom assessment measure (Elhai, Gray, Kashdan, & Franklin, 2005) that has been previously examined for detecting simulated PTSD (Edens, Otto, & Dwyer, 1998; Elhai, Gray, Naifeh, et al., 2005).

EXPERIMENT 1

Method

Participants

A total of 330 students (115 men, 215 women) enrolled in undergraduate psychology courses at two medium-sized state universities in the Midwestern and Western United States participated in this study. We recruited participants from 2003 to 2004 from their classes or departmental research pools in exchange for research or extra course credit. We obtained institutional review board approval for this study.

We conducted the experiment across several semesters in group settings (usually between 20-40 participants). We randomly placed each group of participants into one of four groups in a 2×2 design (thus randomization was at a group level). There were two experimental levels of caution (caution or no caution) and two levels of financial incentive (high incentive and low incentive). Thus, participants constituted one of the following four groups: (a) caution/high incentive ($n = 92$), (b) caution/low incentive ($n = 84$), (c) no caution/high incentive ($n = 88$), and (d) no caution/low incentive ($n = 66$). The sample was predominantly White (92%). Most reported working at least part-time (60%) or were unemployed (40%). The majority was single (91%) or married (7%). Age ranged from 18 to 56 years ($M = 21.13$, $SD = 4.31$), whereas educational level ranged from 11 to 18 years ($M = 13.42$, $SD = 1.40$).

Measures

Life Events Checklist. The Life Events Checklist (LEC; Blake et al., 1995) is a PTSD Criterion A1 measure of exposure to 16 potentially traumatic events. Respondents rate their experience of the event on a 5-point nominal scale (1 = *happened to me;* 2 = *witnessed it;* 3 = *learned about it;* 4 = *not sure;* and 5 = *does not apply*). Test-retest reliability

has been demonstrated, with a mean kappa coefficient of .61 (indicating high correspondence) for direct exposure items and a lower (but acceptable) mean kappa of .41 for indirect exposure items (Gray, Litz, Wang, & Lombardo, 2004). The LEC's items also demonstrated convergent validity with the Traumatic Life Events Questionnaire (Kubany et al., 2000; mean kappa coefficient was .51) and predicted PTSD symptomatology using the PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993) and Modified PTSD Symptom Scale (Falsetti, Resnick, Resick, & Kilkpatrick, 1993, June) in both college students and clinical trauma survivors (Gray et al., 2004).

PCL. The PCL (Weathers et al., 1993) is a 17-item Likert scaled self-report PTSD measure. Respondents rate how much they have been bothered by each symptom in the past month on a 5-point scale (1 = *Not at all;* 2 = *A little bit;* 3 = *Moderately;* 4 = *Quite a bit;* and 5 = *Extremely*) yielding a continuous raw score. The PCL has demonstrated excellent internal consistency ($\alpha = .94$; E. B. Blanchard, Jones-Alexander, Buckley, & Forneris, 1996; Ruggiero, Del Ben, Scotti, & Rabalais, 2003) and test-retest reliability of .88 for a 1-week interval (Ruggiero et al., 2003). E. B. Blanchard et al. (1996) demonstrated sensitivity (.94), specificity (.86), and overall PTSD diagnostic efficiency (.90) using a cutoff score of 44 in motor vehicle accident victims, validated in college students as well (Ruggiero et al., 2003). In the Weathers et al. (1993) study, the PCL correlated positively with the Mississippi PTSD Scale (Keane, Caddell, & Taylor, 1988; .93), MMPI–2 Keane PTSD Scale (Keane, Malloy, & Fairbank, 1984; .77), and Impact of Event Scale (Horowitz, Wilner, & Alvarez, 1979; .90).

TSI. The TSI is a 100-item, standardized posttraumatic symptom measure including 10 clinical and 3 validity scales. Items are rated using a 4-point scale ranging from 0 (*never*) to 3 (*often*) for their frequency over the past 6 months. Raw scores are converted to age- and gender-adjusted standardized T scores. The TSI's Atypical Response scale (ATR) is a fake bad Validity scale consisting of infrequently endorsed items. The TSI scales possess adequate internal consistency ranging from .84 to .87 (Briere, Elliott, Harris, & Cotman, 1995). Concerning validity, it has been shown to share large relationships between its Clinical Scales and established PTSD measures, and 86% to 96% efficiency for regression models using the Clinical and Validity Scales to predict standardized PTSD diagnoses (Briere, 1995; McDevitt-Murphy et al., 2005).

Procedure

We told participants that the overall study's purpose dealt with detecting fabricated PTSD. We administered the following forms, printed with precoded subject numbers (unless otherwise noted), under standard instructions, to groups of participants:

1. Informed consent form (not coded).
2. Contact information form.
3. Demographic survey.
4. LEC (Blake et al., 1995).
5. A brief query of trauma-related characteristics (e.g., age at onset) for endorsed traumatic events including a PTSD Criterion A2 assessment of fear, helplessness, and horror.
6. PCL (Weathers et al., 1993).

Students were then given PTSD training materials and a worksheet to prepare for a subsequent quiz. The training materials included clear descriptions of the 17 PTSD symptoms that are available to the lay person and were adapted from the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) and popular abnormal psychology textbooks. We presented participants with several case studies, and we instructed them to choose only one case study to use.

To investigate the effects of different incentive levels, we randomly told half of the participant groups that the three best PTSD simulators per group of 30 participants would receive $50, $45, or $40 (high incentive condition). We told the other half of groups that the prizes were $15, $10, or $5 (low incentive condition). In both conditions, we emphasized the 1 in 10 odds of winning a prize.

After 25 min of review, experimenters employed by the first author removed the materials and administered a 10-item true-false quiz based on the materials' content. The experimenter then collected and scored all quizzes and informed participants that they passed (to avoid potential embarrassment for those who failed) and could complete the TSI while simulating PTSD. However, it was determined that data would not be used from participants scoring less than 70% (missing more than 3 items) on the quiz.

Next, we administered TSI test booklets and answer sheets preprinted with research numbers. We provided participants verbal and written simulation directions regarding the TSI, adapted from several sources (Bagby et al., 1997; Elhai et al., 2001; Sivec, Hilsenroth, & Lynn, 1995). We instructed participants to simulate PTSD in an attempt to deceive a psychological examiner, with the goal of either obtaining money (personal injury lawsuit or disability payments) or avoiding criminal prosecution because of their PTSD.[1]

To investigate the effects of providing cautionary instructions to simulators, we randomly assigned half of the

[1]We offered participants a choice in the case studies and malingering scenarios used for simulation, which could potentially result in an increased sense of responsibility and investment in simulation. Despite the potential advantage of increased motivation and use in previous studies, we acknowledge that concerns have been raised about this method's generalizability (Rogers, 1997b).

TABLE 1
Experiment 1—Participants' Feedback Response Comparisons

	Simulation Groups									
	Caution/High Incentive n = 91		*Caution/Low Incentive n = 84*		*No Caution/High Incentive n = 87*		*No Caution/Low Incentive n = 65*		*ANOVA*	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F(df)*	η_p^2
Clarity of training	4.86	0.93	5.09	0.80	5.06	.86	4.75	1.13	2.31 (3,323)	.02
Confidence in portraying PTSD	3.88	1.27	3.80	1.02	4.11	1.06	3.80	1.35	1.35 (3, 323)	.01
Extent cash inventive motivated	4.11**	1.48	3.37**	1.64	3.68	1.63	3.46	1.60	3.71 (3, 323)*	.03
Understanding of PTSD	2.79	1.28	3.15	1.46	2.98	1.41	2.82	1.35	1.24 (3, 323)	.01

Note. ANOVA = analysis of variance; PTSD = posttraumatic stress disorder. Effect size convention for η_p^2 = small (.01), medium (.06), and large (.14).
$^*p < .05$. **Tukey's honestly significant difference post hoc test indicated these two means were significantly different at $p < .05$.

participant groups the following additional simulation instructions (caution condition) adapted from Viglione et al. (2001):

> However, please also realize that if you present too dramatically, it will look fake, and you will not be believed. Therefore, use caution to maintain a convincing, realistic profile.

The remaining participants did not receive such instructions (no caution condition).

Next, we asked participants to select which malingering scenario (obtaining money or avoiding prosecution for PTSD) they intended to use when feigning PTSD. After completing the TSI, we collected test materials, and we administered a 6-point Likert scale ranging from 1 (*not at all*) to 6 (*very*) feedback survey, primarily measuring perceptions of effort, motivation from the cash incentive, and training material adequacy.

We determined prize winners from participants whose total summed score on the TSI's Intrusive Experiences, Defensive Avoidance, and Anxious Arousal scales (corresponding to the three PTSD symptom criteria) most closely resembled that of a PTSD patient group's mean score (Davis et al., 2006) without elevating ATR above a T score of 90 (Briere, 1995). We awarded prizes by mail.

Results

Feedback Response Comparisons

We examined participants' perceptions of their training, effort, and motivation. Groups did not score significantly different regarding perceived clarity of training, confidence in portraying PTSD, or understanding of PTSD (see Table 1). Results were significant for participants' report of motivation by the cash incentive; however, the effect size was small. Results of a post hoc comparison using Tukey's honestly significant difference test ($p < .05$) revealed that participants in the caution/high incentive group reported being significantly more motivated by the cash incentive to successfully feign PTSD than participants in the caution/low incentive group (see Table 1 for means and standard deviations). Finally, the type of case study chosen for simulation did not differ by simulation group, $\chi^2(3, N = 328) = .69$, $p > .05$.

Data Exclusions

We screened participants for possible PTSD and then excluded them from the primary analyses. *Possible PTSD* is a term we derived to indicate that participants met the following set of experiment-related PTSD criteria (which, albeit, does not include a clinical interview): (a) reported experiencing or witnessing a traumatic event(s) on the LEC, (b) endorsed PTSD's Criterion A2, and (c) met or exceeded the established PCL cutoff score of 44 (Ruggiero et al., 2003). This procedure resulted in the removal of 51 participants from the primary analyses (17 from the caution/high incentive group, 9 from the caution/low incentive group, 17 from the no caution/high incentive group, and 8 from the no caution/low incentive group). Additionally, we removed one simulator for scoring less than 70% on the PTSD quiz, and we removed 20 simulators for scoring 75 or higher on the TSI's Inconsistency scale (INC) or for leaving 10 or more items blank. These exclusions resulted in a final sample of 70 caution/high incentive, 70 caution/low incentive, 68 no caution/high incentive, and 50 no caution/low incentive participants.

Primary Results

We utilized two-way analyses of variance (ANOVAs) to assess for both main effects of caution and incentive and interactions on ATR and the 10 Clinical Scales. We adjusted the significance levels for these analyses to .01 to adjust for multiple comparisons.

Main effects of caution were significant for ATR and all Clinical Scales except for the Dysfunctional Sexual Behavior and Tension Reduction scales (see Table 2). However, a main effect of incentive was only significant for the Impaired Self-Reference scale. Further, caution and incentive interacted only on the Dissociation scale.

TABLE 2
Experiment 1—Simulation Groups' Descriptive Statistics and ANOVA Results for the Trauma Symptom Inventory (TSI)

	Simulation Groups								*Main Effects*					
	Caution/High Incentive[a]		*Caution/Low Incentive*[b]		*No Caution/High Incentive*[c]		*No Caution/Low Incentive*[d]		*Caution*		*Incentive*		*Interaction*	
TSI Scale	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F(1,254)*	*Cohen's d*	*F(1,254)*	*Cohen's d*	*F(1,254)*	η_p^2
Response Atypical	66.27	17.34	69.60	16.93	73.87	17.79	74.54	17.23	8.27*	.36	0.84	.11	0.37	.00
Arousal Anxious	62.90	8.97	65.00	7.80	69.25	9.03	68.33	9.15	19.49*	.55	0.29	.07	1.88	.01
Depression	61.01	8.63	63.00	8.02	65.14	8.68	65.72	9.21	10.03*	.40	1.41	.15	0.42	.00
Anger/Irritability	59.74	9.07	63.11	10.22	65.83	10.11	64.83	11.05	9.52*	.39	.88	.12	2.99	.01
Intrusive Experiences	70.03	9.80	71.67	8.25	76.91	8.86	77.68	8.15	33.63*	.72	1.17	.14	0.15	.00
Defensive Avoidance	63.35	6.82	66.11	6.04	68.47	7.52	67.56	7.72	13.93*	.47	1.11	.13	4.35	.02
Dissociation	62.48	10.34	68.67	10.30	71.47	12.62	68.62	11.45	10.06*	.40	1.41	.15	10.33*	.04
Sexual Concerns	52.94	10.07	56.16	10.46	57.07	11.35	59.50	13.67	6.94*	.33	3.97	.25	0.08	.00
Dysfunctional Sexual Behavior	55.85	14.25	58.58	14.45	58.75	15.85	62.37	18.45	2.91	.21	2.61	.20	0.05	.00
Impaired Self-Reference	57.75	9.27	61.54	7.98	62.40	9.13	64.53	9.11	11.75*	.43	7.04*	.33	0.55	.00
Tension Reduction Behavior	63.79	12.59	67.02	13.75	69.55	14.33	69.98	14.97	6.26	.31	1.11	.13	0.65	.00

Note. ANOVA = analysis of variance. Effect size cnventions for Cohen's d = small (.20), medium (.50), and large (.80). Effect size conventions for η_p^2 = small (.01), medium (.06), and large (.14).
[a]$n = 70$. [b]$n = 70$. [c]$n = 68.0$ [d]$n = 50$.
*$p < .01$.

EXPERIMENT 2

Method

Participants

A total of 180 adults (59 men, 121 women) participated in this study. Participants were college students from the same universities who were recruited, randomized, and participated in the experiment in the same manner as in Experiment 1. We also obtained institutional review board approval for this study. We randomized experimental groups to either the PTSD ($n = 88$) or PTSD-MDD ($n = 92$) conditions. Most participants (95%) were White. The majority reported part-time (58%) or no (37%) employment. Relationship status for the majority (81%) was single. Age ranged from 18 to 48 years, with a mean of 21.74 years ($SD = 5.12$). Years of education ranged from 12 to 17 and averaged 13.64 ($SD = 1.33$).

Measures

The measures in Experiment 2 were identical to the measures utilized in Experiment 1.

Procedure

The procedure utilized in Experiment 2 was identical to the procedure of Experiment 1, with the following exceptions. Participants in the PTSD-MDD group received additional training materials on MDD (similar in design to the PTSD training materials all participants received), and we instructed them that successful PTSD simulation also requires MDD feigning. When participants in the PTSD-MDD group completed the TSI, we instructed them to simulate PTSD (as were participants in the PTSD group), but we reminded them to also simulate MDD.

We informed all participants in this experiment that the three best PTSD simulators per group of 30 participants would receive $50, $45, or $40 (respectively) and emphasized the 1 in 10 odds of winning a prize. Furthermore, we cautioned all to appear believable in their simulation. Finally, we gave participants in the PTSD-MDD group a four-item, true-false MDD quiz (to test their knowledge of MDD) in addition to the 10-item true-false PTSD quiz.[2] We excluded participants who scored less than 75% (missed more than one item) on the four-item MDD quiz from further analyses due to a presumed lack of effort.

Results

Feedback Response Comparisons

We examined perceptions of study training, effort, and motivation. Groups did not score significantly different regarding perceived clarity of training, report of motivation by the cash incentive, or confidence in portraying PTSD. However, PTSD-MDD participants reported a greater prior

[2]Both the 10-item PTSD and 4-item MDD quizzes are available for review to interested researchers. Please contact J. D. Elhai to obtain these materials.

TABLE 3
Experiment 2—Participants' Feedback Response Comparisons

	Simulation Groupd							
	PTSD			*PTSD–MDD*			*ANOVA*	
Variable	*n*	*M*	*SD*	*n*	*M*	*SD*	*F(df)*	*Cohen's d*
Clarity of training	87	4.84	0.99	92	5.04	0.82	2.13 (1, 177)	.22
Confidence in portraying PTSD	87	4.05	1.17	92	4.18	1.02	.72 (1, 177)	.13
Extent incentive cash motivated	86	3.98	1.41	90	3.88	1.56	.19 (1, 177)	.07
Understanding of PTSD	86	3.06	1.22	90	3.40	0.99	4.04 (1, 174)*	.30

Note. PTSD = posttraumatic stress isorder; MDD = major depressive disorder; ANOVA = analysis of variance. Effect size conventions for Cohen's d = small (.20), medium (.50), and large (.80).
$^{*}p < .05$.

understanding of PTSD (see Table 3), but this effect was small. Finally, groups did not score significantly different regarding the case study chosen, $\chi^2(1, N = 178) = 1.83$, $p > .05$.

Data Exclusions

We screened participants for possible PTSD and then excluded them from the primary analyses, which resulted in the removal of 42 participants (21 from each group). Additionally, we removed four simulators for scoring less than 70% on the PTSD or 75% on the MDD quizzes due to presumed lack of effort. Based on suggested TSI exclusion criteria (Briere, 1995), we also removed six simulators for scoring 75 or higher on the INC scale or for leaving 10 or more TSI items blank. These exclusions resulted in a remaining sample of 60 PTSD and 68 PTSD-MDD simulators (128 participants).

Primary Results

We compared groups on ATR and the 10 Clinical Scales using one-way ANOVAs. Using an alpha level of .01 (to adjust for multiple comparisons), we found no significant differences between simulation groups on any scales (see Table 4).

TABLE 4
Experiment 2—Simulation Groups' Descriptive Statistics and ANOVA Results for the Trauma Symptom Inventory (TSI)

	Simulation Groups					
	PTSD		*PTSD–MDD*			
TSI Scale	*M*	*SD*	*M*	*SD*	*F(1,127)**	*Cohen's d*
Atypical Response	67.77	17.02	68.66	14.96	0.10	.03
Anxious Arousal	64.61	8.04	66.87	7.80	2.59	.29
Depression	62.28	9.56	65.06	8.32	3.09	.31
Anger/Irritability	59.36	9.92	60.74	9.77	0.63	.14
Intrusive Experiences	71.01	9.58	71.80	8.97	0.23	.08
Defensive Avoidance	64.46	7.89	65.22	5.73	0.39	.11
Dissociation	66.01	10.24	68.45	9.04	2.04	.25
Sexual Concerns	55.57	10.88	56.00	11.10	0.05	.04
Dysfunctional Sexual Behavior	57.97	14.45	57.06	14.98	0.12	.06
Impaired Self-Reference	60.65	9.18	61.38	8.58	0.21	.08
Tension Reduction Behavior	63.74	14.46	67.10	14.21	2.01	.25

*None significant at $p < .01$.
Note. ANOVA = analysis of variance; PTSD = posttraumatic stress disorder; MDD = major depressive disorder. Effect size conventions for Cohen's d small (.20), Medium (.50), and large (.80)

GENERAL DISCUSSION

In Experiment 1, we examined effects and interactions of cautionary statements and financial incentive levels on simulated PTSD. We hypothesized that cautionary instructions would produce large effects, but different financial incentive levels would produce small or no effects.

Hypotheses regarding caution main effects were supported, as cautioned participants appeared significantly more conservative in their symptom endorsement on ATR and all TSI Clinical Scales except for Dysfunctional Sexual Behavior and Tension Reduction Behavior, with effect sizes ranging from small to medium. Our hypotheses concerning the Intrusive Experiences, Defensive Avoidance, and Anxious Arousal scales were also supported. These three TSI Clinical Scales had the largest effect sizes (.72, .47, and .55, respectively) of any significant TSI scale. This is intuitively consistent with our methodology of training participants to simulate PTSD specifically and with the fact that these scales are most closely related to the three primary PTSD symptom criteria. These findings support those of previous studies, demonstrating that simple cautionary instructions regarding simulation believability can be effective in altering clinical presentations (Viglione et al., 2001). Further, these findings provide important confirmation that malingering researchers should consider the implications of using cautionary instructions and the potential impact on feigning effort.

Concerning incentive main effects, only the Impaired Self-Reference Clinical scale was significant. This result supports some previous studies that have demonstrated a lack of effect for different incentives used (Bernard, 1990; Martin et al., 1993). These findings potentially suggest that the use

of larger incentives (e.g., $50) over smaller ones (e.g., $15) in malingering research may not be necessary (at least not when simulating psychopathology on briefer instruments such as the TSI). However, given the lack of research in this area, caution should be exercised in drawing strong conclusions. For example, we note that this study is the first that assessed differential incentive levels, and true differences may exist that are simply difficult to detect with the small incentive amounts typically offered in analogue studies. Additionally, it is still unclear to what extent college or general population participants will be motivated by these amounts. Previous research has suggested that small financial amounts may be meaningful to those with few financial resources, such as incarcerated inmates, but not to others (Schretlen, Wilkins, Van Gorp, & Bobholz, 1992). Additionally, financial incentives may conceivably have a greater impact in real-world settings when the feigner has a greater opportunity to prepare to malinger as compared to typical analogue studies in which participants often have just a few minutes for preparation. Further research is clearly warranted in this area.

Interaction effects were nonsignificant with the exception of the Dissociation scale. Although an interesting finding, it is difficult to explain an interaction of caution and incentive on this scale alone. As previously discussed, the Dissociation scale did yield a main effect for caution but not for incentive (consistent with most other TSI scales). The Dissociation scale is thought to measure dissociative symptomatology such as depersonalization, out-of-body experiences, and psychic numbing (Briere, 1995). Future research will help to clarify whether a particular sensitivity of the combined effects of caution and financial incentives to dissociation simulation potentially exists.

Finally, the results of Experiment 1 that indicate that caution/high incentive participants reported more motivation by the cash incentive than caution/low incentive participants was a similarly interesting yet also unexpected finding. Although the effect size was small, it potentially suggests a caution and incentive interaction that could be explored in future research. However, this finding only represents participants' report of motivation and does not represent actual scale differences, and thus should be interpreted cautiously. Further research will add important data to this question.

Our aim in Experiment 2 was to determine if PTSD simulators trained in a commonly occurring comorbid diagnosis (MDD) would yield an altered PTSD simulation performance than without such comorbidity training. Results demonstrate that comorbid training did not affect simulators' performances, with no statistically different TSI scales. This finding has the potential to suggest that alterations in common malingering research methodology regarding comorbidity may not be necessary. However, two notes of caution are important to mention before drawing firm conclusions. First, virtually no previous research has been conducted in this area, and we are thus unable to provide directional interpretations. Second, and most important, the interesting findings of Experiment 1 suggest that another possible explanation for the lack of significant findings may have been a suppressive effect of the cautionary statement used in this experiment. Experiment 1 findings, consistent with Viglione et al. (2001), demonstrate that cautionary instructions regarding simulation believability resulted in participants who appeared more conservative in symptom endorsement. The cautionary statement used with all participants in Experiment 2 may have contributed to the lack of significant findings. These possibilities provide a clear call for further research and provide exciting opportunities to clarify these questions. For example, future studies might explore comorbid diagnostic training with varying conditions of cautionary statement use. Also, future investigators may employ a more general psychopathology instrument (e.g., MMPI–2) with other forms of simulated psychopathology.

Finally, our hypothesis regarding an elevation of the TSI's Depression scale in the PTSD-MDD group was also not supported. This finding was surprising given that the Depression scale was specifically designed to assess for depressed mood and cognitions (Briere, 1995). A suppressive effect of the cautionary statement again becomes a potential consideration. Another possible factor in explaining this finding was whether the p value ($<.01$) utilized was overly stringent. However, the TSI Depression scale score (actual p value $=$ .08) suggests this was not the case.

In summary, both experiments yielded interesting results that have significant future research implications. Experiment 1 demonstrated that cautionary instructions may result in more conservative symptom endorsement. Results concerning the use of financial incentives indicate that this practice continues to be a less well-understood research methodological procedure that clearly calls for further study. Also, although this study failed to provide strong evidence of interactions between these two procedures, there is some evidence it should be studied further. Finally, Experiment 2 failed to demonstrate clear suggestions for comorbid diagnostic training, yet the possibility that cautionary statements were instrumental in the lack of findings also suggest further research is warranted.

Several limitations were inherent in this study. First, the simulation sample was a convenience sample comprised solely of college students and may not be representative of community participants. Second, the methodology included only PTSD simulation (rather than other disorders) and pertained only to the TSI, and thus, generalizability of the results cannot be confirmed. Nonetheless, the study's strengths help contribute to the growing literature examining malingering research designs and provides exciting ideas for future research.

ACKNOWLEDGMENTS

This study was based on the master's theses of J. J. Butcher, A. N. Reeves, and S. N. Baugher and was funded by the

Office of Research and Graduate Education, The University of South Dakota. Simulation training materials are available by request from J. D. Elhai.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bagby, R. M., Buis, T., & Nicholson, R. A. (1995). Relative effectiveness of the standard validity scales in detecting fake-bad and fake-good responding: Replication and extension. *Psychological Assessment, 7,* 84–92.

Bagby, R. M., Rogers, R., & Buis, T. (1994). Detecting malingered and defensive responding on the MMPI-2 in a forensic inpatient sample. *Journal of Personality Assessment, 62,* 191–203.

Bagby, R. M., Rogers, R., Nicholson, R., Buis, T., Seeman, M. V., & Rector, N. (1997). Does clinical training facilitate feigning schizophrenia on the MMPI-2? *Psychological Assessment, 9,* 106–112.

Bernard, L. C. (1990). Prospects for faking believable memory deficits on neuropsychological tests and the use of incentives in simulation research. *Journal of Clinical and Experimental Neuropsychology, 12,* 715–728.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., et al. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress, 8,* 75–90.

Blanchard, D. D., McGrath, R. E., Pogge, D. L., & Khadivi, A. (2003). A comparison of the PAI and MMPI-2 as predictors of faking bad in college students. *Journal of Personality Assessment, 80,* 197–205.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behavior Research and Therapy, 34,* 669–673.

Briere, J. (1995). *Trauma Symptom Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.

Briere, J., Elliott, D. M., Harris, K., & Cotman, A. (1995). Trauma Symptom Inventory: Psychometrics and association with childhood and adult victimization in clinical samples. *Journal of Interpersonal Violence, 10,* 387–401.

Bury, A. S., & Bagby, R. M. (2002). The detection of feigned uncoached and coached posttraumatic stress disorder with the MMPI-2 in a sample of workplace accident victims. *Psychological Assessment, 14,* 472–484.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI–2: Minnesota Multiphasic Personality Inventory–2: Manual for administration and scoring.* Minneapolis: University of Minnesota Press.

Davis, J. L., Borntrager, C. F., Combs-Lane, A. M., Wright, D. C., Elhai, J. D., Falsetti, S. A., et al. (2006). Comparison of racial groups on trauma and post-trauma functioning. *Journal of Trauma Practice, 5,* 21–36.

Edens, J. F., Otto, R. K., & Dwyer, T. J. (1998). Susceptibility of the Trauma Symptom Inventory to malingering. *Journal of Personality Assessment, 71,* 379–392.

Elhai, J. D., Gold, P. B., Frueh, B. C., & Gold, S. N. (2000). Cross-validation of the MMPI-2 in detecting malingered posttraumatic stress disorder. *Journal of Personality Assessment, 75,* 449–463.

Elhai, J. D., Gold, S. N., Sellers, A. H., & Dorfman, W. I. (2001). The detection of malingered posttraumatic stress disorder with MMPI–2 fake bad indices. *Assessment, 8,* 221–236.

Elhai, J. D., Gray, M. J., Kashdan, T. B., & Franklin, C. L. (2005). Which instruments are most commonly used to assess traumatic event exposure and posttraumatic effects?: A survey of traumatic stress professionals. *Journal of Traumatic Stress, 18,* 541–545.

Elhai, J. D., Gray, M. J., Naifeh, J. A., Butcher, J. J., Davis, J. L., Falsetti, S. A., et al. (2005). Utility of the Trauma Symptom Inventory's Atypical Response Scale in detecting malingered post-traumatic stress disorder. *Assessment, 12,* 210–219.

Elhai, J. D., Naifeh, J. A., Zucker, I. S., Gold, S. N., Deitsch, S. E., & Frueh, B. C. (2004). Discriminating malingered from genuine civilian posttraumatic stress disorder: A validation of three MMPI-2 Infrequency Scales (F, Fp, and Fptsd). *Assessment, 11,* 139–144.

Elhai, J. D., Ruggiero, K. J., Frueh, B. C., Beckham, J. C., Gold, P. B., & Feldman, M. E. (2002). The Infrequency-Posttraumatic Stress Disorder scale (Fptsd) for the MMPI–2: Development and initial validation with veterans presenting with combat-related PTSD. *Journal of Personality Assessment, 79,* 531–549.

Falsetti, S. A., Resnick, H. S., Resick, P. A., & Kilkpatrick, D. G. (1993, June). The Modified PTSD Symptom Scale: A brief self-report measure of posttraumatic stress disorder. *Behavior Therapist, 16,* 161–162.

Franklin, C. L., & Zimmerman, M. (2001). Posttraumatic stress disorder and major depressive disorder: Investigating the role of overlapping symptoms in diagnostic comorbidity. *Journal of Nervous and Mental Disease, 189,* 548–551.

Frueh, B. C., Hamner, M. B., Cahill, S. P., Gold, P. B., & Hamlin, K. (2000). Apparent symptom overreporting among combat veterans evaluated for PTSD. *Clinical Psychology Review, 20,* 853–885.

Gray, M. J., Litz, B. T., Wang, J., & Lombardo, T. W. (2004). Psychometric properties of the Life Events Checklist. *Assessment, 11,* 330–341.

Guriel, J., & Fremouw, W. (2003). Assessing malingered posttraumatic stress disorder: A critical review. *Clinical Psychology Review, 23,* 881–904.

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of event scale: A measure of subjective stress. *Psychosomatic Medicine, 41,* 209–218.

Keane, T. M., Caddell, J. M., & Taylor, K. L. (1988). Mississippi Scale for Combat-Related Posttraumatic Stress Disorder: Three studies in reliability and validity. *Journal of Consulting and Clinical Psychology, 56,* 85–90.

Keane, T. M., Malloy, P., & Fairbank, J. (1984). Empirical development of an MMPI subscale for the assessment of combat-related posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 52,* 888–891.

Keane, T. M., & Wolfe, J. (1990). Comorbidity in post-traumatic stress disorder: An analysis of community and clinical studies. *Journal of Applied Social Psychology, 20,* 1776–1788.

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry, 52,* 1048–1060.

Knoll, J. L., & Resnick, P. J. (1999). U.S. v. Greer: Longer sentences for malingerers. *Journal of the American Academy of Psychiatry and Law, 27,* 621–625.

Kubany, E. S., Haynes, S. N., Leisen, M. B., Owens, J. A., Kaplan, A. S., Watson, S. B., et al. (2000). Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: The Traumatic Life Events Questionnaire. *Psychological Assessment, 12,* 210–224.

Martin, R. C., Bolter, J. F., Todd, M. E., Gouvier, W. D., & Niccolls, R. (1993). Effects of sophistication and motivation on the detection of malingered memory performance using a computerized forced-choice task. *Journal of Clinical and Experimental Neuropsychology, 15,* 867–880.

McDevitt-Murphy, M. E., Weathers, F. W., & Adkins, J. W. (2005). The use of the Trauma Symptom Inventory in the assessment of PTSD symptoms. *Journal of Traumatic Stress, 18,* 63–67.

McNally, R. J. (2003). Progress and controversy in the study of posttraumatic stress disorder. *Annual Review of Psychology, 54,* 229–252.

Niccols, R., & Bolter, J. (1991). *Multi-digit Memory Test*. San Luis Obisto, CA: Wang Neuropsychological Laboratories.

Nicholson, R. A., Mouton, G. J., Bagby, R. M., Buis, T., Peterson, S. A., & Buigas, R. A. (1997). Utility of MMPI-2 indicators of response distortion: Receiver operating characteristic analysis. *Psychological Assessment, 9,* 471–479.

Resnick, P. J. (1997). Malingering of posttraumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 130–152). New York: Guilford.

Rey, A. (1941). L'examen psychologigique dans le cas d'encephalopathie traumatique. *Archives de Psychologie, 37,* 126–139.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Rogers, R. (1997a). Introduction. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 1–19). New York: Guilford.

Rogers, R. (1997b). Researching dissimulation. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 398–426). New York: Guilford.

Rogers, R., & Cruise, K. R. (1998). Assessment of malingering with simulation designs: Threats to external validity. *Law and Human Behavior, 22,* 273–285.

Ruggiero, K. J., Del Ben, K., Scotti, J. R., & Rabalais, A. E. (2003). Psychometric properties of the PTSD Checklist-Civilian Version. *Journal of Traumatic Stress, 16,* 495–502.

Schretlen, D., Wilkins, S. S., Van Gorp, W. G., & Bobholz, J. H. (1992). Cross-validation of a psychological test battery to detect faked insanity. *Psychological Assessment, 4,* 77–83.

Sivec, H. J., Hilsenroth, M. J., & Lynn, S. J. (1995). Impact of simulating borderline personality disorder on the MMPI-2: A costs-benefits model employing base rates. *Journal of Personality Assessment, 64,* 295–311.

Smith, G. P., & Burger, G. K. (1997). Detection of malingering: Validation of the Structured Inventory of Malingered Symptomatology (SIMS). *Journal of the Academy of Psychiatry and Law, 25,* 183–189.

Viglione, D. J., Wright, D. M., Dizon, N. T., Moynihan, J. E., DuPuis, S., & Pizitz, T. D. (2001). Evading detection on the MMPI-2: Does caution produce more realistic patterns of responding? *Assessment, 8,* 237–250.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993, October). *The PTSD checklist: Reliability, validity, & diagnostic utility.* Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Weschler, D. (1987). *Weschler Memory Scale-Revised*. New York: The Psychological Corporation.

Jon D. Elhai
Disaster Mental Health Institute
The University of South Dakota
414 East Clark Street-SDU 114
Vermillion, SD 57069-2390
Email: jonelhai@fastmail.fm or jelhai@usd.edu

Received June 16, 2005
Revised July 24, 2006