Psychological Medicine (2009), **39**, 1957–1966.
doi:10.1017/S0033291709005819

ORIGINAL ARTICLE

Diagnostic alterations for post-traumatic stress disorder: examining data from the National Comorbidity Survey Replication and National Survey of Adolescents

Jon D. Elhai[1]*, Julian D. Ford[2], Kenneth J. Ruggiero[3] and B. Christopher Frueh[4]

[1] *University of South Dakota, Vermillion, SD, USA*
[2] *University of Connecticut Medical School, Farmington, CT, USA*
[3] *Medical University of South Carolina, Charleston, SC, USA*
[4] *Baylor College of Medicine and the Menninger Clinic, Houston, TX, USA*

Background. Two alternative models of post-traumatic stress disorder (PTSD) appear to represent the disorder's latent structure better than the traditional Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) three-factor PTSD model. The present study examines the impact of using these structural models for the diagnosis of lifetime PTSD while retaining the DSM-IV PTSD's six-symptom diagnostic requirement.

Method. Data were gathered from large-scale, epidemiological datasets collected with adults (National Comorbidity Survey Replication) and adolescents (National Survey of Adolescents). Two alternative, empirically supported four-factor models of PTSD were compared with the DSM-IV three-factor PTSD diagnostic model.

Results. Results indicated that the diagnostic alterations resulted in substantially improved structural validity, downward adjustments of PTSD's lifetime prevalence (roughly 1 percentage point decreases in adults, 1–2.5 percentage point decreases in adolescents), and equivalent psychiatric co-morbidity and sociodemographic associations.

Conclusions. Implications for modifying PTSD diagnostic criteria in future editions of DSM are discussed.

Received 13 October 2008; Revised 9 March 2009; Accepted 14 March 2009; First published online 20 April 2009

Key words: Construct validity, diagnosis, DSM, epidemiology, post-traumatic stress disorder, trauma.

Introduction

The diagnostic model of post-traumatic stress disorder (PTSD) currently includes three symptom clusters, labeled 're-experiencing' (criterion B), 'avoidance and numbing' (criterion C) and 'hyperarousal' (criterion D) (APA, 2001). However, a substantial body of empirical research has demonstrated that this model does not adequately explain the structure of the PTSD construct (Asmundson *et al*. 2004; Frueh *et al*. 2004). Furthermore, alternative models have not been tested for their convergence in diagnosing Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) PTSD.

Numerous recent studies have used confirmatory factor analysis (CFA) to test specific, theoretically driven models of PTSD in revealing the hypothesized model(s) best fitting, or accounting for, respondents' intercorrelation patterns (for a review, see Asmundson *et al*. 2004). This literature, with various trauma-exposed samples, has consistently found that the traditional three-factor DSM-IV PTSD model does not adequately fit observed data (e.g. DuHamel *et al*. 2004; McWilliams *et al*. 2005; Naifeh *et al*. 2008) and that other models fit better (e.g. Elhai *et al*. 2007, 2008; Elkit & Shevlin, 2007; Palmieri *et al*. 2007*b*; Naifeh *et al*. 2008; Saul *et al*. 2008).

Two models in particular have garnered the most empirical support for explaining PTSD's factor structure. Although other models have been investigated, the two models highlighted here consistently statistically outperform them (most recently in Elkit & Shevlin, 2007; Krause *et al*. 2007; Naifeh *et al*. 2008; Palmieri *et al*. 2007*a*, *b*; Saul *et al*. 2008).

The revised DSM-IV PTSD model of King *et al*. (1998) separates criterion C's effortful avoidance and emotional numbing symptoms into separate factors,

* Address for correspondence: J. D. Elhai, Ph.D., Disaster Mental Health Institute, The University of South Dakota, 414 East Clark Street – SDU 114, Vermillion, South Dakota 57069-2390, USA.
(Email: jonelhai@gmail.com)

resulting in re-experiencing (B1–B5), effortful avoidance (C1–C2), emotional numbing (C3–C7) and hyperarousal (D1–D5). This model reflects research demonstrating that avoidance and numbing are statistically distinct constructs, revealing divergent patterns of correlates with psychopathology, and different prognoses and treatment effects (for a review, see Asmundson *et al*. 2004). Numerous CFA studies have supported the King model with adults (most recently in Elhai *et al*. 2007, 2008; Palmieri *et al*. 2007*a*; Schinka *et al*. 2007) and adolescents (Saul *et al*. 2008).

The model of Simms *et al*. (2002) retains the King model's re-experiencing (B1–B5) and effortful avoidance (C1–C2) factors. However, three PTSD hyperarousal symptoms – sleep difficulty (D1), irritability (D2) and concentration problems (D3) – are moved from the hyperarousal factor and combined with the emotional numbing factor's symptoms (C3–C7) to form a 'dysphoria' factor. Thus, the resulting model includes re-experiencing (B1–B5), effortful avoidance (C1–C2), dysphoria (C3–C7 and D1–D3) and hyperarousal (D4–D5). The Simms model reflects empirical work demonstrating that general distress/dysphoria is an underlying component of mood and anxiety disorders (including PTSD) (e.g. Watson, 2005; Slade & Watson, 2006), and that this construct is distinct from the hyperarousal associated specifically with post-traumatic reactions (Simms *et al*. 2002). Several recent studies have supported this model over other similar models with adults (Elkit & Shevlin, 2007; Krause *et al*. 2007; Palmieri *et al*. 2007*b*), but it has not been examined with adolescents.

At present, the King and Simms models appear to be the top candidates for best explaining PTSD's symptom structure. However, no studies have tested the practical impact of altering PTSD's diagnosis based on the alternative King or Simms models. And because the PTSD factor analytic literature illuminates how symptoms empirically cluster together, drawing from that literature can guide the psychiatry field toward improved PTSD diagnostic models.

The present investigation, using large-scale, epidemiological data, examines the impact of altering the PTSD diagnosis based on these models, testing differences in PTSD's prevalence, co-morbidity and sociodemographics, and factor structure. We explored our hypotheses in both samples of adolescents and adults, to test our research questions across two separate age groups. We hypothesized that the diagnostic alterations would result in the following: (1) a downward adjustment of PTSD's prevalence, due to more conservatively stringent PTSD diagnostic criteria; (2) no substantial decreases in co-morbidity rates, because recent research found that even removing PTSD symptoms that overlap with highly co-morbid mood/anxiety disorders did not affect PTSD's co-morbidity (Elhai *et al*. 2008); and (3) significant improvements in PTSD's factor structure, based on the CFA findings for PTSD summarized above. We retained the required number of six symptoms in altering the diagnosis, in order to reliably compare results with the same six-symptom requirement currently used in the DSM-IV PTSD diagnosis.

This study is important and timely since the PTSD construct has received criticism recently for a range of conceptual issues pertaining to symptom structure and symptom overlap with other mental disorders (Spitzer *et al*. 2007; Rosen & Lilienfield, 2008; Rosen *et al*. 2008). Also, decreasing PTSD's symptom overlap with mood and anxiety disorders would ensure that the PTSD construct is unique as a diagnostic entity. Further, conceptual issues in the mental disorders (including PTSD) are currently being considered for DSM-V, and investigations such as this one potentially speak to the issue of whether and how to revise the disorder's diagnostic criteria.

Study 1: method

Sample

Study 1 used archival data from the National Comorbidity Survey Replication (NCS-R) (Kessler, 2006). The NCS-R was a nationally stratified, multistage area household probability sample study of non-institutionalized adults (aged 15 years and older). The NCS-R was conducted with 9282 participants in the early 2000s (NCS-R part I), with demographic characteristics presented in previous NCS-R reports (Kessler *et al*. 2004). The present paper reports on the representative subsample of participants completing the NCS-R part II (which evaluated PTSD; $n = 5692$).

Instruments

The World Mental Health Survey Initiative version of the structured Composite International Diagnostic Interview (CIDI; Kessler & Ustun, 2004) was used to diagnose DSM-IV mental disorders; the CIDI evidences adequate convergence with other similar measures (Andrews & Peters, 1998; Haro *et al*. 2006). In addition to the lifetime DSM-IV PTSD diagnostic variable, we also examined lifetime diagnostic variables for mood, anxiety and substance-use disorders. DSM-IV diagnostic algorithms were used to assign diagnoses, discussed elsewhere (Kessler *et al*. 2005). We exclusively used the NCS-R's 'non-hierarchy' diagnoses (i.e. allowing a particular diagnosis to be assigned even if it occurred solely in the presence of another disorder).

Of relevance to PTSD, participants were first asked in behaviorally specific terms about previous exposure to a variety of traumatic events meeting DSM-IV's PTSD stressor criterion (A1). Only those participants endorsing a traumatic event with initial fear, helplessness or horror (criterion A2) were subsequently queried about DSM-IV PTSD symptoms. PTSD symptom queries involved binary ('yes'/'no') lifetime symptom ratings about one's trauma. For those endorsing more than one trauma, the most upsetting occurrence of their most upsetting traumatic event type was used. For individuals whose most upsetting trauma occurrence was different from a trauma that was randomly selected by NCS-R investigators, they were instructed to rate their PTSD symptoms separately for each event; PTSD diagnoses were then assigned based on meeting PTSD criteria from either of these two events. Finally, PTSD's criteria E (duration) and F (functional impairment) were queried. Skip-out rules were implemented, such that if a participant did not meet a particular PTSD symptom criterion, s/he was not subsequently queried about remaining PTSD criteria (discussed further below).

Analyses

NCS-R part II sampling weights were used for all analyses in study 1, to adjust for differential household size, non-response and post-stratification. We used Stata 9.0 software (StataCorp LP, College Station, TX, USA) to examine whether the King and Simms models yielded differences from the DSM-IV PTSD model in prevalence of PTSD, diagnostic status, diagnostic co-morbidity and sociodemographic associations. We used Mplus 5.1 software (Muthén & Muthén, 1998–2007) to examine the impact of using these models on PTSD's structural validity.

We calculated the lifetime DSM-IV PTSD diagnosis, and King and Simms alterations using the NCS-R's PTSD module item data, retaining DSM-IV PTSD's six-symptom diagnostic requirement. For the PTSD diagnosis (and alterations), we required that the respondent endorsed at least one traumatic event meeting criteria A1 and A2, and satisfied criteria E and F. For criteria B through D in the traditional DSM-IV PTSD diagnosis, we required at least one re-experiencing symptom (B1–B5), at least three avoidance/numbing symptoms (C1–C7) and at least two hyperarousal symptoms (D1–D5) (as required in DSM-IV). To remain consistent with the six-item diagnostic requirement in DSM-IV, for the King PTSD diagnostic alteration we retained the requirements for at least one re-experiencing symptom (B1–B5) and at least two hyperarousal symptoms (D1–D5); in addition, we required at least one avoidance symptom (C1–C2) and at least two numbing symptoms (C3–C7) (thus for the similar DSM-IV PTSD minimum total of three avoidance/numbing symptoms). For the Simms PTSD diagnostic alteration, we retained the requirement for at least one re-experiencing symptom (B1–B5) and the King model's requirement of at least one avoidance symptom (C1–C2); with only two possible hyperarousal symptoms in this model, we required at least one hyperarousal symptom (D4–D5), leaving a requirement of at least three dysphoria symptoms (C3–C7, D1–D3) in order to yield a total of six required PTSD symptoms.

For structural validity analyses, we only used data from participants who (in addition to endorsing PTSD's criterion A1) had endorsed criterion A2 in reference to an index trauma on which PTSD ratings were made ($n=871$). For participants with multiple sets of PTSD ratings, we used ratings from their most upsetting event. We further excluded 105 subjects missing more than four (24%) of PTSD's symptom items (leaving 766 remaining participants) because of skip-out rules; the remaining participants represented a slightly skewed sample in that they met at least one PTSD symptom cluster (only 42 subjects were skipped out of two symptom clusters). Additional missing item-level data (typically, one or two items each, by 6% of subjects) were estimated to preserve the sample size, using maximum likelihood (ML) estimation of missing data (for a review, see Schafer & Graham, 2002) for categorical outcomes (Muthén & Muthén, 1998–2007, p. 401).

Since the observed dependent variables were binary ('yes'/'no') ratings, we implemented robust (mean- and variance-adjusted) weighted least squares (WLS) estimation for the CFAs, using polychoric (rather than Pearson) correlations and probit (rather than linear) regression coefficients (Flora & Curran, 2004; Wirth & Edwards, 2007). In fitting the CFAs, we estimated intercorrelations among the common factors, not allowing intercorrelations among the residuals' error variances. χ^2 Tests of model fit were examined in conjunction with goodness-of-fit indices, including the Tucker–Lewis index (TLI), comparative fit index (CFI) and root mean square error of approximation (RMSEA) (interpreted when RMSEA <0.06 for an excellent fit, and between 0.06 to 0.08 for an adequate fit; CFI/TLI $\geqslant 0.95$ for an excellent fit, and between 0.90 to 0.94 for an adequate fit) (Hu & Bentler, 1998, 1999). Differences between the DSM-IV and King PTSD models were tested using a χ^2 difference test (albeit with a correction factor since a robust χ^2 statistic was used; Muthén & Muthén, 2006). When comparing the Simms model with the DSM-IV and King PTSD models, these models are not nested subsets within each other, and thus χ^2 difference testing is not

Table 1. *PTSD prevalence across the DSM-IV, King and Simms PTSD diagnostic models*

	DSM-IV model		King model		Simms model	
Sample	*n*	% (S.E.)	*n*	% (S.E.)	*n*	% (S.E.)
NCS-R	387.53	6.81 (0.36)	339.07	5.96 (0.03)[a]	332.83	5.85 (0.03)[b,c]
NSA	329	8.18 (0.43)	277	6.89 (0.40)[d]	226	5.61 (0.36)[e,f]

PTSD, Post-traumatic stress disorder; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; S.E., standard error; NCS-R, National Comorbidity Survey Replication; NSA, National Survey of Adolescents.

[a] Significantly different prevalence from the NCS-R DSM-IV model (binomial approximation $z=2.55$, S.E. $=0.00$, $p=0.02$).

[b] Significantly different prevalence from the NCS-R DSM-IV model (binomial approximation $z=2.88$, S.E. $=0.00$, $p<0.01$).

[c] Not significantly different prevalence from the NCS-R King model (binomial approximation $z=0.35$, S.E. $=0.00$, $p>0.05$).

[d] Significantly different prevalence from the NSA DSM-IV model (binomial approximation $z=2.99$, S.E. $=0.00$, $p<0.01$).

[e] Significantly different prevalence from the NSA DSM-IV model (binomial approximation $z=5.95$, S.E. $=0.00$, $p<0.001$).

[f] Significantly different prevalence from the NSA King model (binomial approximation $z=3.21$, S.E. $=0.00$, $p=0.002$).

Table 2. *Changes in PTSD status between DSM-IV diagnostic system and King and Simms diagnostic alterations*

	King alteration PTSD diagnostic determination, *n*[a]		Simms alteration PTSD diagnostic determination, *n*[a]	
DSM-IV PTSD diagnostic determination	Not diagnosed	Diagnosed	Not diagnosed	Diagnosed
National Comorbidity Survey Replication Data				
Not diagnosed	5305	0	5280	25
Diagnosed	48	339	79	308
National Survey of Adolescents Data				
Not diagnosed	3694	0	3669	25
Diagnosed	52	277	128	201

PTSD, Post-traumatic stress disorder; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition.

[a] Weighted numbers, rounded to the nearest whole number.

appropriate; instead, differences in Bayesian information criterion (BIC) values were examined. BIC values, only estimable using ML (but not WLS) estimation, were generated using ML-estimated CFAs, with logistic rather than probit coefficients [more appropriate using ML (Muthén, 1984) albeit without using sampling weights which cannot be currently implemented with ML for categorical outcomes]. When a model has a BIC value of 10 points less than another model, there is a 150:1 odds that the model with the smaller BIC value is the better-fitting model (Raftery, 1995).

Study 1: results

Prevalence

Table 1 illustrates PTSD prevalence rates across the DSM-IV, King and Simms models, with binomial approximation *z* tests presented to statistically compare proportions (Hays, 1994). The King and Simms alterations each reduced DSM-IV PTSD's prevalence by nearly one percentage point.

Table 2 demonstrates changes in DSM-IV PTSD status with the King diagnostic alteration. The level of agreement between these diagnostic systems could be considered 'almost perfect' (based on criteria set forth by Landis & Koch, 1977) [$\kappa = 0.92$, standard error (S.E.) = 0.01 (98.63% agreement), $z = 69.79$, $p < 0.001$]. Only 0.84% of the entire sample (48 weighted subjects out of 5692 subjects) changed diagnostic status for PTSD when using the King criteria.

The level of agreement between the DSM-IV and Simms models was 'almost perfect' (Landis and Koch, 1977) [$\kappa = 0.84$, S.E. = 0.01 (97.24% agreement), $z = 63.82$, $p < 0.001$]. In combination, approximately 1.82% of the entire sample (79 + 25 weighted subjects out of 5692 subjects) changed diagnostic status for PTSD when using the Simms criteria. Additionally, the King and Simms diagnostic systems were highly concordant with each other (considered 'almost perfect' agreement) [$\kappa = 0.86$, S.E. = 0.01 (97.73% agreement), $z = 65.05$, $p < 0.001$].

Finally, we assessed if particular PTSD symptom clusters were responsible for the decreased PTSD prevalence estimates seen in the King and Simms diagnostic alterations. Among respondents meeting DSM-IV PTSD criteria A, E and F, we estimated the proportion meeting only one particular DSM-IV PTSD symptom criterion (among re-experiencing; avoidance/numbing; and arousal) compared with the proportion meeting only the one comparable DSM-IV cluster for the King model (re-experiencing; avoidance and numbing; and arousal) and Simms model (re-experiencing; avoidance and dysphoria; and arousal). The King model's decreased PTSD prevalence resulted solely from its avoidance and numbing symptom diagnostic alteration, and the Simms model's prevalence decrease resulted solely from its arousal symptom alteration.

Co-morbidity and sociodemographic variables

Next we examined if the PTSD diagnostic alterations were associated with different prevalence estimates of co-morbid mental disorders and sociodemographic characteristics (Table 3). Regardless of whether the diagnostic alterations were implemented, similar rates of diagnostic co-morbidity and sociodemographic associations were found.

Structural validity

The three-factor DSM-IV PTSD model did not fit the data well [robust $\chi^2(31, n = 766) = 94.76$, $p < 0.001$, TLI = 0.89, CFI = 0.89, RMSEA = 0.05, BIC = 14638.73]. Most problematic, standardized factor loadings and R^2 values were quite low for avoiding thoughts/feelings of the trauma (criterion C1: $\beta = 0.25$, $R^2 = 0.06$), with numerous unstandardized residual covariances high (> 0.10; Kline, 2004), and these parameters were not as problematic for the King (C1: $\beta = 0.44$, $R^2 = 0.19$) or Simms (C1: $\beta = 0.45$, $R^2 = 0.20$) models. The King model had an adequate (but not excellent) fit [robust $\chi^2(31, n = 766) = 86.14$, $p < 0.001$, TLI = 0.91, CFI = 0.90, RMSEA = 0.05, BIC = 14619.07], and this model was significantly better than the three-factor DSM-IV model ($p < 0.001$). Finally, the Simms model fit the data very well [robust $\chi^2(31, n = 766) = 55.90$, $p < 0.001$, TLI = 0.96, CFI = 0.96, RMSEA = 0.03, BIC = 14533.31]. These results support data from two recent studies examining PTSD's symptom structure using the NCS-R data (Cox *et al*. 2008, Elhai *et al*. 2008). However, those studies did not test the Simms model (goodness-of-fit indices here differ slightly from the study of Cox *et al*., which used different statistical software that generates less conservative fit indices).

Finally, one interpretation of the Simms model is that the dysphoria symptom cluster should be de-emphasized from the PTSD diagnosis (Simms *et al*. 2002). Therefore, we tested a three-factor alternative to the Simms model that simply removes the dysphoria factor, which fit the data quite well [robust $\chi^2(11, n = 766) = 20.33$, $p = 0.04$, TLI = 0.95, CFI = 0.97, RMSEA = 0.03, BIC = 7724.45].

In comparing models, judging from BIC values, the original Simms model appears to have approximately a 1600:1 chance of being the preferred model over the three-factor DSM-IV model, and a 1300:1 chance of being the preferred model over the King model; the Simms model that removed the dysphoria symptoms had the best-fitting BIC value.

Study 2: method

Sample

Study 2 used archival data from the National Survey of Adolescents (NSA; Kilpatrick & Saunders, 1995, 1997), a cross-sectional, nationally representative sample of 4023 adolescents aged 12 to 17 years. Demographic characteristics are reported in previous NSA papers (e.g. Kilpatrick *et al*. 2000).

Instruments

Family and youth demographic characteristics were obtained. Structured diagnostic interview questions in the NSA assessed each criterion item from the DSM-IV for lifetime major depressive episode and substance-use disorders (for specific details on instrumentation, see Appendix of Kilpatrick *et al*. 2000). Exposure to potentially traumatic events (PTSD's criterion A1) was assessed using behaviorally specific terms. No direct assessment of PTSD's criterion A2 was conducted.

Lifetime PTSD symptoms (criteria B through D) were assessed using binary 'yes'/'no' ratings based

Table 3. *Diagnostic co-morbidity and sociodemographic differences between DSM-IV, King and Simms PTSD diagnoses in the National Comorbidity Survey Replication data*

	DSM-IV PTSD diagnosis		King PTSD diagnosis			Simms PTSD diagnosis		
	% (S.E.)	n^{a}	% (S.E.)	n^{a}	z^{b}	% (S.E.)	n^{a}	z^{c}
Specific co-morbid disorder diagnosed[d]								
MDE	54.72 (2.73)	212	54.34 (2.97)	184	0.15	53.36 (2.96)	178	0.52
GAD	28.03 (2.15)	108	28.73 (2.35)	97	0.30	28.19 (2.33)	94	0.07
Specific phobia	35.21 (2.42)	136	35.70 (2.62)	121	0.20	37.29 (2.65)	124	0.83
Social phobia	33.39 (2.30)	129	34.67 (2.52)	118	0.52	35.66 (2.56)	119	0.92
Alcohol abuse/dependence	28.22 (2.33)	109	27.53 (2.46)	93	0.29	28.95 (2.51)	96	0.31
Sociodemographic variable								
Female gender	75.41 (2.28)	292	75.93 (2.46)	257	0.23	75.97 (2.52)	253	0.25
Age 40 years or higher	55.12 (2.67)	213	54.07 (2.92)	183	1.70	54.10 (3.05)	180	0.39
White, non-Hispanic	72.60 (2.38)	281	71.57 (2.60)	243	0.44	72.54 (2.60)	241	0.03
At least some college education	51.79 (2.71)	200	52.87 (2.94)	179	0.41	49.70 (3.12)	165	0.79
Employed	65.93 (2.50)	255	67.04 (2.68)	227	0.44	64.99 (2.82)	216	0.38
Currently married	50.51 (2.73)	196	50.46 (2.97)	171	0.02	49.88 (3.11)	166	0.24
Recent mental healthcare use[e]	38.01 (2.53)	147	39.00 (2.77)	132	0.39	40.47 (2.93)	135	0.96
PTSD onset post age 17 years	53.79 (2.76)	208	52.27 (3.00)	177	0.58	50.81 (3.12)	169	1.14
Disability[f]	23.76 (2.05)	92	24.89 (2.26)	84	0.51	25.00 (2.28)	83	0.55

DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; PTSD, post-traumatic stress disorder; S.E., standard error; MDE, major depressive disorder; GAD, generalized anxiety disorder.

[a] Weighted numbers, rounded to the nearest whole number.

[b] Binomial approximation z test statistic for proportions, comparing co-morbidity rates between the DSM-IV and King PTSD diagnostic systems, using an average sample size of 364 across diagnostic systems.

[c] Binomial approximation z test statistic for proportions, comparing co-morbidity rates between the DSM-IV and Simms PTSD diagnostic systems, using an average sample size of 361 across diagnostic systems.

[d] Only the most prevalent co-morbid disorders in this sample are displayed here.

[e] Recent mental healthcare use was defined as at least one treatment visit in the previous year.

[f] Disability involved if the respondent cut down or was unable to perform usual activities due to a mental health problem for at least 1 day in the previous month.

on symptoms that occurred at any time in the adolescent's life, lasting at least 2 weeks, but were not linked to an index trauma. Participants were also queried about PTSD's criterion F. No skip-out rules were used. The PTSD interview was adapted from a similar interview that demonstrated adequate diagnostic convergence with other measures (Kilpatrick *et al*. 1998).

Analyses

We used the same statistical software and analytic strategy presented in study 1, above.

We first calculated the DSM-IV PTSD diagnosis, and King and Simms alterations using the NSA's PTSD module item data. For diagnosing PTSD, we required that the respondent endorsed at least one previous trauma meeting DSM-IV PTSD's criterion A1 (items on witnessing violent trauma; sexual assault or physical assault or abuse; and direct exposure to disaster, accident, actual or threatened serious injury)[1]†. Consistent with the DSM-IV, we also required functional impairment for a PTSD diagnosis. For criteria B through D, we implemented the same requirements as stipulated in study 1, above.

For structural validity analyses, only data from the 3351 trauma-exposed participants were used. Less than 3% of participants were missing data (typically one or two items each), estimated as discussed above in study 1.

Study 2: results

Prevalence

Table 1 illustrates PTSD prevalence across the DSM-IV, King and Simms PTSD models, compared using binomial approximation z tests. The diagnostic

† The notes appear at the end of the main text.

Table 4. *Diagnostic co-morbidity differences between DSM-IV, King and Simms PTSD diagnoses in the National Survey of Adolescents data*

	DSM-IV PTSD diagnosis		King PTSD diagnosis			Simms PTSD diagnosis		
	% (S.E.)	n^a	% (S.E.)	n^a	z^b	% (S.E.)	n^a	z^c
Specific disorders diagnosed								
MDE	75.68 (2.37)	249	76.90 (2.54)	213	0.42	75.66 (2.86)	171	0.01
Alcohol abuse	13.68 (1.90)	45	14.44 (2.12)	40	0.38	15.49 (2.41)	35	0.88
Drug abuse	5.78 (1.29)	19	5.05 (1.32)	14	0.54	7.08 (1.71)	16	0.93
Sociodemographic variable								
Female gender	68.69 (3.18)	226	69.61 (3.43)	193	0.35	67.97 (3.78)	154	0.26
Age 15–17 years	67.76 (3.20)	223	67.40 (3.49)	187	0.13	62.75 (3.92)	142	1.79
White, non-Hispanic	67.76 (3.20)	223	67.96 (3.48)	188	0.07	67.97 (3.78)	154	0.07
Ninth-grade education or higher	77.57 (2.86)	255	77.90 (3.09)	216	0.14	72.55 (3.62)	164	1.17
PTSD onset post age 13 years	50.00 (3.43)	165	50.28 (3.73)	139	0.10	45.75 (4.04)	103	1.42

DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, 4th edition; PTSD, post-traumatic stress disorder; S.E., standard error; MDE, major depressive disorder.

[a] Unweighted numbers, rounded to the nearest whole number.

[b] Binomial approximation z test statistic for proportions, comparing co-morbidity rates between the DSM-IV and King PTSD diagnostic systems, using an average sample size of 303 across diagnostic systems.

[c] Binomial approximation z test statistic for proportions, comparing co-morbidity rates between the DSM-IV and Simms PTSD diagnostic systems, using an average sample size of 278 across diagnostic systems.

alterations resulted in roughly a 1–2.5 percentage point drop in DSM-IV PTSD's prevalence.

Table 2 demonstrates changes in DSM-IV PTSD status with the King diagnostic alteration. The level of agreement between these diagnostic systems was 'almost perfect' (Landis & Koch, 1977) [$\kappa = 0.91$, S.E. = 0.02 (98.71% agreement), $z = 57.79$, $p < 0.001$]. Only 1.29% of the entire sample (52 out of 4023 subjects) changed diagnostic status for PTSD when using the King criteria.

Table 2 also demonstrates changes in DSM-IV PTSD status with the Simms diagnostic alteration. The level of agreement between these diagnostic systems was 'substantial' (Landis & Koch, 1977) [$\kappa = 0.70$, S.E. = 0.02 (96.20% agreement), $z = 45.60$, $p < 0.001$]. In combination, approximately 3.80% of the entire sample (128 + 25 weighted subjects out of 4023 subjects) changed diagnostic status for PTSD when using the Simms criteria. Also, the King and Simms diagnostic systems were highly concordant with each other and considered to represent 'substantial' agreement (Landis & Koch, 1977) [$\kappa = 0.71$, S.E. = 0.02 (96.69% agreement), $z = 45.82$, $p < 0.001$].

Finally, we assessed if particular PTSD symptom clusters were responsible for the decreased PTSD prevalence estimates seen in the King and Simms diagnostic alterations, as conducted in study 1. Again, the King model's decreased PTSD prevalence resulted solely from its avoidance and numbing symptom diagnostic alteration, and the Simms model's prevalence decrease resulted solely from its arousal symptom alteration.

Co-morbidity and sociodemographic variables

Next we examined if the PTSD diagnostic alterations were associated with different prevalence rates of co-morbid depression and substance-use disorders (but not drug dependence, whereby cell sizes were too small) and sociodemographic characteristics. Table 4 illustrates that regardless of whether the diagnostic alterations were implemented, extremely similar rates of diagnostic co-morbidity and sociodemographic associations were found.

Structural validity

The three-factor DSM-IV PTSD model fit the data very well [robust $\chi^2(99, n = 3351) = 343.37$, $p < 0.001$, TLI = 0.99, CFI = 0.98, RMSEA = 0.03, BIC = 40647.69]. The King model also fit the data very well [robust $\chi^2(97, n = 3351) = 268.55$, $p < 0.001$, TLI = 0.99, CFI = 0.98, RMSEA = 0.02, BIC = 40590.65], and this model was significantly better than the three-factor DSM-IV model ($p < 0.001$). The Simms model fit the data very well [robust $\chi^2(97, n = 3351) = 245.69$, $p < 0.001$, TLI = 0.99, CFI = 0.99, RMSEA = 0.02, BIC = 40573.27]. Finally, we tested a three-factor Simms model that removes the dysphoria factor, which fit the data quite well [robust $\chi^2(22,\ n=3351)=80.34,\ p<0.001$, TLI = 0.99, CFI = 0.99, RMSEA = 0.03, BIC = 20390.34].

Judging from BIC values, the Simms model appears to have approximately a 1120:1 chance of being the preferred model over the three-factor model, and a 260:1 chance of being the preferred model over the King model; however, the model removing the dysphoria factor appears best-fitting.

General discussion

Results of CFAs demonstrate that both the King and Simms models represent significant improvements over the current DSM-IV PTSD model. The Simms model resulted in the factor structure with the best fit to the data, followed by the King model; this finding was opposite from that found in a similar paper using the original NCS data (McWilliams *et al*. 2005), possibly due to the more sophisticated methods of querying trauma and PTSD in the NCS-R. These findings are important because data-reductionistic studies of the PTSD construct have been lacking (Frueh *et al*. 2004; Buckley, 2007) and it is not clear whether the current DSM-IV PTSD model is efficient, parsimonious, theoretically consistent, or empirically supported in contrast to other similar models. Given that both alternative models provided such significant improvements over the DSM-IV model (and in previous studies), it seems likely that meaningful changes to the diagnostic criteria may reasonably be expected in future revisions of the DSM.

Lower PTSD prevalence estimates in the King or Simms diagnostic alterations could be due to several reasons: (1) 'under-detection' of PTSD cases in that some symptomatic/impaired individuals may be missed as true PTSD cases; (2) increased diagnostic precision in correctly diagnosing impaired individuals and correctly ruling out non-impaired individuals; and (3) (consistent with supplementary analyses), the separation of PTSD's active avoidance from emotional numbing (King model), as well as restricting hyperarousal symptoms to the truly post-traumatic hyperarousal that does not involve dysphoria (Simms model). Given that both the King and Simms models outperformed the DSM-IV model in CFAs, it seems plausible to accept that the 'true' PTSD prevalence may be slightly lower than previously understood.

The King model resulted in the downward adjustment of DSM-IV PTSD's prevalence by roughly one percentage point (Table 1) because 12–16% of DSM-IV PTSD cases were not classified with PTSD using the King model (Table 2). Thus the King model would represent a more conservative PTSD diagnosis than the current DSM-IV's PTSD diagnosis, resulting in better specificity in ruling out non-PTSD cases, but fewer impaired trauma victims meeting the diagnosis of PTSD. However, given the importance of effortful avoidance and its clear distinction from emotional numbing (for a review, see Asmundson *et al*. 2004), adding such stringency may more accurately describe the psychopathology of the PTSD construct than the DSM-IV PTSD model currently does.

In contrast, the Simms model converged less with the traditional DSM-IV PTSD model. After all, the Simms model not only separates the DSM-IV PTSD model's avoidance from numbing symptoms, but also separates purely post-traumatic hyperarousal symptoms from dysphoria symptoms. Compared with DSM-IV PTSD's prevalence, the Simms model's overall prevalence was lower by only one percentage point in the NCS-R, but by 2.5 points in the NSA (Table 1). Moreover, the Simms model classified 20% (NCS-R) to 40% (NSA) of the DSM-IV PTSD cases as non-cases (Table 2). It is true that requiring pure hyperarousal as distinct from dysphoria appears to have substantial empirical support (Simms *et al*. 2002; Krause *et al*. 2007; Palmieri *et al*. 2007*b*). Thus the Simms diagnostic alteration may be best in quite conservatively diagnosing PTSD, but leads to a question of whether too many legitimate PTSD cases may be ruled out.

The DSM-IV PTSD model is the 'gold standard' for diagnosis. However, it may in fact simply reflect an arbitrary *status quo* for diagnosis. The Simms model may be best at capturing the PTSD construct's underlying symptom structure, based on CFA findings. However, in terms of practical diagnosis of PTSD, the King model may serve as a compromise between accurately reflecting the disorder's distinct key components while not overhauling the diagnosis or substantially reducing its prevalence. This emphasis on diagnosis is not meant to suggest that PTSD should be viewed strictly as a dichotomous 'present'/'absent' diagnostic phenomenon, since the taxometric literature suggests that PTSD is a dimensional disorder (e.g. Broman-Fulks *et al*. 2006). However, assigned diagnoses are still used and required in healthcare practice settings, and for communicating among treatment providers.

Although large-scale, epidemiological datasets were used in the present paper, several limitations should nonetheless be considered. For example, we retained the required six symptoms needed for the PTSD diagnosis, and it is quite possible that relaxing or tightening this requirement would result in different findings. However, varying the minimum symptom requirement was not the focus of this paper; rather, we aimed to test whether other established models with the same required symptom counts would perform differently from the DSM-IV PTSD model.

Second, PTSD symptom queries in the NSA were not linked to an index traumatic event, which could have led to inflated PTSD prevalence (Long & Elhai, in press).

Despite its limitations, this research contributes to our understanding of the PTSD construct's symptom structure and how empirically supported PTSD factor models can translate to meaningful changes in diagnosing the disorder. The increase in specificity that would come with the King and Simms models should be placed in the context of the PTSD diagnosis' trend toward less stringency over time, and such liberalizing of the diagnosis has been criticized (McNally, 2003; Long & Elhai, in press). As deliberation about psychiatric nosology continues in the next several years in preparation for the next (fifth) edition of the DSM, these and other complementary findings concerning the structure and clinical utility of the PTSD diagnosis will be important in informing decisions that will affect tens of thousands of traumatized people for whom valid assessment and diagnosis are crucial to the receipt of evidence-based treatment and recovery from post-traumatic impairment.

Acknowledgements

Dr Frueh is now affiliated with the University of Hawaii-Hilo. Jon D. Elhai, Ph.D., is now affiliated with the University of Toledo.

Declaration of Interest

None.

Note

[1] Kilpatrick *et al*. (2000) did not use direct exposure to disaster, accident, or actual or threatened serious injury to diagnose PTSD, but the present report did, to remain consistent with DSM-IV PTSD's criterion A1. Other potentially relevant traumatic events queried in the NSA were not used to indicate trauma exposure (in Kilpatrick *et al*. 2000 nor in the present report), including: serious injury, illness or death of a family member or close friend; and major personal illness or injury to oneself. These events were defined with such ambiguity that it is difficult to assess if they met PTSD's criterion A1, and because they used a 1-year (rather than lifetime) time-frame.

References

Andrews G, Peters L (1998). The psychometric properties of the Composite International Diagnostic Interview. *Social Psychiatry and Psychiatric Epidemiology* **33**, 80–88.

APA (2001). *Diagnostic and Statistical Manual of Mental Disorders*, 4th edn, text revision. American Psychological Association: Washington, DC.

Asmundson GJG, Stapleton JA, Taylor S (2004). Are avoidance and numbing distinct PTSD symptom clusters? *Journal of Traumatic Stress* **17**, 467–475.

Broman-Fulks JJ, Ruggiero KJ, Green BA, Kilpatrick DG, Danielson CK, Resnick HS, Saunders BE (2006). Taxometric investigation of PTSD: Data from two nationally representative samples. *Behavior Therapy* **37**, 364–380.

Buckley TC (2007). PTSD and Vietnam veterans. *Science* **315**, 185.

Cox BJ, Mota N, Clara I, Asmundson GJG (2008). The symptom structure of posttraumatic stress disorder in the National Comorbidity Replication Survey. *Journal of Anxiety Disorders* **22**, 1523–1528.

DuHamel KN, Ostrof J, Ashman T, Winkel G, Mundy EA, Keane TM, Morasco BJ, Vickberg SMJ, Hurley K, Burkhalter J, Chhabra R, Scigliano E, Papadopoulos E, Moskowitz C, Redd W (2004). Construct validity of the Posttraumatic Stress Disorder Checklist in cancer survivors: analyses based on two samples. *Psychological Assessment* **16**, 255–266.

Elhai JD, Gray MJ, Docherty AR, Kashdan TB, Kose S (2007). Structural validity of the Posttraumatic Stress Disorder Checklist among college students with a trauma history. *Journal of Interpersonal Violence* **22**, 1471–1478.

Elhai JD, Grubaugh AL, Kashdan TB, Frueh BC (2008). Empirical examination of a proposed refinement to DSM-IV posttraumatic stress disorder symptom criteria using the National Comorbidity Survey Replication data. *Journal of Clinical Psychiatry* **69**, 597–602.

Elkit A, Shevlin M (2007). The structure of PTSD symptoms: A test of alternative models using confirmatory factor analysis. *British Journal of Clinical Psychology* **46**, 299–313.

Flora DB, Curran PJ (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* **9**, 466–491.

Frueh BC, Elhai JD, Kaloupek DG (2004). Unresolved issues in the assessment of trauma exposure and posttraumatic reactions. In *Posttraumatic Stress Disorder: Issues and Controversies* (ed. G. M. Rosen), pp. 63–84. Wiley: New York.

Haro JM, Arbabzadeh-Bouchez S, Brugha TS, De Girolamo G, Guyer ME, Jin R, Lepine JP, Mazzi F, Reneses B, Vilagut G, Sampson NA, Kessler RC (2006). Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health Surveys. *International Journal of Methods in Psychiatric Research* **15**, 167–180.

Hays W (1994). *Statistics*. Wadsworth: Belmont, CA.

Hu L, Bentler PM (1998). Fit indices in covariance structural modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods* **3**, 424–453.

Hu L, Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling* **6**, 1–55.

Kessler RC (2006). National Comorbidity Survey: Replication (NCS-R), 2001–2003 (computer file). Conducted by Harvard Medical School, Department of Health Care Policy/University of Michigan, Survey Research Center. ICPSR04438-v1. Inter-University

Consortium for Political and Social Research: Ann Arbor, MI.
Kessler RC, Berglund P, Chiu WT, Demler O, Heeringa S, Hiripi E, Jin R, Pennell BE, Walters EE, Zaslavsky A, Zheng H (2004). The US National Comorbidity Survey Replication (NCS-R): design and field procedures. *International Journal of Methods in Psychiatric Research* **13**, 69–92.
Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry* **62**, 593–602.
Kessler RC, Ustun TB (2004). The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research* **13**, 93–121.
Kilpatrick DG, Acierno R, Saunders B, Resnick HS, Best CL, Schnurr PP (2000). Risk factors for adolescent substance abuse and dependence: data from a national sample. *Journal of Consulting and Clinical Psychology* **68**, 19–30.
Kilpatrick DG, Resnick HS, Freedy JR, Pelcovitz D, Resick P, Roth S, van der Kolk B (1998). Posttraumatic stress disorder field trial: evaluation of the PTSD construct – criteria A through E. In *DSM-IV Sourcebook* (ed. T. A. Widiger, A. J. Frances, H. A. Pincus, R. Ross, M. B. First, W. Davis and M. Kline), pp. 804–844. American Psychiatric Press: Washington, DC.
Kilpatrick DG, Saunders BE (1995). National Survey of Adolescents in the United States (computer file). ICPSR version. Inter-university Consortium for Political and Social Research: Ann Arbor, MI.
Kilpatrick DG, Saunders BE (1997). Prevalence and consequences of child victimization. Results from the National Survey of Adolescents: Final report. National Crime Victims Research and Treatment Center: Charleston, SC.
King DW, Leskin GA, King LA, Weathers FW (1998). Confirmatory factor analysis of the Clinician-Administered PTSD Scale: evidence for the dimensionality of posttraumatic stress disorder. *Psychological Assessment* **10**, 90–96.
Kline RB (2004). *Principles and Practice of Structural Equation Modeling*. Guilford Press: New York.
Krause ED, Kaltman S, Goodman LA, Dutton MA (2007). Longitudinal factor structure of posttraumatic stress symptoms related to intimate partner violence. *Psychological Assessment* **19**, 165–175.
Landis JR, Koch GG (1977). The measurement of observer agreement for categorical data. *Biometrics* **33**, 159–174.
Long ME, Elhai JD (in press). Posttraumatic stress disorder's traumatic stressor criterion: History, controversy, clinical and legal implications. *Psychological Injury and Law*.
McNally RJ (2003). Progress and controversy in the study of posttraumatic stress disorder. *Annual Review of Psychology* **54**, 229–252.
McWilliams LA, Cox BJ, Asmundson GJG (2005). Symptom structure of posttraumatic stress disorder in a nationally representative sample. *Journal of Anxiety Disorders* **19**, 626–641.
Muthén BO (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika* **49**, 115–132.
Muthén BO, Muthén L (2006). *Chi-square Difference Testing Using the Satorra-Bentler Scaled Chi-square*. Muthén & Muthén: Los Angeles.
Muthén LK, Muthén BO (1998–2007). *Mplus User's Guide*. Muthén & Muthén: Los Angeles.
Naifeh JA, Elhai JD, Kashdan TB, Grubaugh AL (2008). The PTSD Symptom Scale's latent structure: an examination of trauma-exposed medical patients. *Journal of Anxiety Disorders* **22**, 1355–1368.
Palmieri PA, Marshall GN, Schell TL (2007*a*). Confirmatory factor analysis of posttraumatic stress symptoms in Cambodian refugees. *Journal of Traumatic Stress* **20**, 207–216.
Palmieri PA, Weathers FW, Difede J, King DW (2007*b*). Confirmatory factor analysis of the PTSD Checklist and the Clinician-Administered PTSD Scale in disaster workers exposed to the World Trade Center Ground Zero. *Journal of Abnormal Psychology* **116**, 329–341.
Raftery AE (1995). Bayesian model selection in social research. *Sociological Methodology* **25**, 111–163.
Rosen GM, Lilienfield SO (2008). Posttraumatic stress disorder: an empirical evaluation of core assumptions. *Clinical Psychology Review* **28**, 837–868.
Rosen GM, Spitzer RL, McHugh PR (2008). Problems with the post-traumatic stress disorder diagnosis and its future in DSM V. *British Journal of Psychiatry* **192**, 3–4.
Saul AL, Grant KE, Carter JS (2008). Post-traumatic reactions in adolescents: how well do the DSM-IV PTSD criteria fit the real life experience of trauma exposed youth? *Journal of Abnormal Child Psychology* **17**, 277–289.
Schafer JL, Graham JW (2002). Missing data: our view of the state of the art. *Psychological Methods* **7**, 147–177.
Schinka JA, Brown LM, Borenstein AR, Mortimer JA (2007). Confirmatory factor analysis of the PTSD Checklist in the elderly. *Journal of Traumatic Stress* **20**, 281–289.
Simms LJ, Watson D, Doebbeling BN (2002). Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology* **111**, 637–647.
Slade T, Watson D (2006). The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. *Psychological Medicine* **36**, 1593–1600.
Spitzer RL, First MB, Wakefield JC (2007). Saving PTSD from itself in DSM-V. *Journal of Anxiety Disorders* **21**, 233–241.
Watson D (2005). Rethinking the mood and anxiety disorders: a quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology* **114**, 522–536.
Wirth RJ, Edwards MC (2007). Item factor analysis: current approaches and future directions. *Psychological Methods* **12**, 58–79.