Addressing the Limitations of Confirmatory Factor Analysis of Posttraumatic Stress Disorder: An Application of Exploratory Structural Equation Modeling

by

Jessica Marie Petri

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements of the Degree of
Doctor of Philosophy

Auburn, Alabama
August 7, 2021

Keywords: Posttraumatic stress disorder; exploratory structural equation modeling; confirmatory factor analysis; construct validity research

Approved by

Frank Weathers, Chair, Professor of Psychology
Tracy Witte, Associate Professor of Psychology
Christopher Correia, Professor of Psychology
Jeffrey Katz, Professor of Psychology

Abstract

To date, confirmatory factor analysis (CFA) has been the predominant method for empirically evaluating the latent factor structure of posttraumatic stress disorder (PTSD). A voluminous CFA literature provided a strong empirical rationale for separating effortful avoidance from emotional numbing symptoms in the *Diagnostic and Statistical Manual of Mental Disorders*, *Fifth Edition* (*DSM-5*; APA, 2013) PTSD criteria. Nevertheless, CFA studies of PTSD also have generated increasingly complex structural models, which lack *a priori* theoretical justification and parsimony, yield highly correlated factors, and include multiple factors with only two indicators. It is possible that the CFA requirement of constraining factor cross-loadings to zero has biased the findings of previous PTSD CFAs. Accordingly, the primary aim of the present study was to evaluate *DSM-5* PTSD factor structure using exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009), which permits cross-loadings and thereby might circumvent some of the limitations of CFA (Prudon, 2015). Participants were trauma-exposed undergraduates (N = 1,139) who completed the PTSD Checklist for *DSM-5* (PCL-5; Weathers et al. 2013a). Analyses focused on comparisons between the four-factor *DSM-5* model of PTSD and the seven-factor Hybrid model, using both ESEM and CFA. As expected, in both four- and seven-factor ESEM models there were a number of significant cross-loadings, which justified the use of ESEM. For both ESEM and CFA, the seven-factor Hybrid model was superior in fit to the four-factor *DSM-5* model, and thus was accepted as the best representation of PTSD. However, compared to its seven-factor CFA counterpart, the seven-factor ESEM model was superior in fit, had lower factor loadings and factor intercorrelations, and exhibited better-differentiated patterns of associations with external correlates. These findings suggest that ESEM is more useful than CFA for investigations of

PTSD factor structure. They also suggest that the superior fit of the seven-factor Hybrid model is not attributable solely to CFA misspecifications, and thus support its continued use in PTSD construct validity research.

Acknowledgments

I would like to acknowledge the expertise and guidance of Dr. Frank Weathers and Dr. Tracy Witte on this research project and throughout my graduate training. I am in debt to the mentorship of Dr. Christian End who was a formative figure in my early pursuits of psychological research. I would also like to thank my lab and cohort members at Auburn University for their understanding and humor through times of stress. Lastly, I am grateful to my partner for providing unconditional support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

α	alpha
AA	Anxious Arousal latent factor
AAR	Alterations in Arousal and Reactivity latent factor
ALC	Alcohol Problems latent factor
AN	Anhedonia latent factor
ANT	Antisocial Features latent factor
ANX	Anxiety latent factor
APA	American Psychiatric Association
AV	Avoidance latent factor
C	Scaling factor
CAPS-5	Clinician Administered PTSD Scale for *DSM-5*
CFA	Confirmatory factor analysis
CFA	Comparative fit index
DA	Dysphoric Arousal latent factor
DEP	Depression latent factor
df	Degrees of freedom
DRG	Drug Problems latent factor
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, 4th Edition
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
EB	Externalizing Behavior latent factor
EFA	Exploratory factor analysis
ESEM	Exploratory structural equation modeling

FIML	Full information maximum likelihood estimation
ICM	Independent clusters model
LEC-5	Life Events Checklist for *DSM-5*
M	Mean
N	Sample size
NA	Negative Affect latent factor
NACM	Negative Alterations in Cognitions and Mood latent factor
p	p value
PAI	Personality Assessment Inventory
PCL-5	PTSD Checklist for *DSM-5*
PTSD	Posttraumatic stress disorder
q	Cohen's q effect size statistic
REX	Reexperiencing latent factor
RMSEA	Root mean square error of approximation
SD	Standard deviation
SRMR	Standardized root mean square residual
TLI	Tucker-Lewis fit index
W	Wald W test statistic
χ^2	Chi-square test statistic

Introduction

Over the last 25 years, posttraumatic stress disorder (PTSD) has been the subject of extensive factor structure investigation (Armour et al., 2016a; Elhai & Palmieri, 2011; Palmieri, Weathers, Difede, & King, 2007). Structural studies of PTSD have relied on factor analysis, an effective and popular tool for empirically identifying and validating theorized dimensions of PTSD (Palmieri et al., 2007). Early studies applied exploratory factor analysis (EFA) to self-reported indicators of PTSD (Taylor, Kuch, Koch, Crockett, & Passey, 1998), largely producing two-factor solutions inconsistent with leading theories of PTSD at the time (Armour et al., 2016a, 2016b; Kilpatrick, 2013; Palmieri et al., 2007; Taylor et al., 1998). However, EFA is not a practical approach for confirming predicted relationships and factor structure because it is applied without an *a priori* model (Palmieri et al., 2007) and often yields atheoretical results (Armour et al., 2016a; Elhai & Palmieri, 2011; Prudon, 2015). EFA models are also strongly affected by extreme item inter-correlations or items with very low or very high prevalence, producing factor solutions that may differ in content and number from the original measure (Prudon, 2015). Lastly, EFA is not suitable for addressing other important questions in latent variable modeling, including tests of invariance, complex error structures, growth modeling, latent mean structures, differential item functioning, and bifactor models (Marsh et al., 2009).

Thus, the majority of studies have relied on confirmatory factor analysis (CFA) to empirically evaluate PTSD factor structure (Armour et al., 2016a, 2016b; Elhai & Palmieri, 2011; Palmieri et al., 2007). When applying CFA, researchers specify *a priori* the relationship between observed indicators (items) and the underlying construct dimensions (latent factors). Consequently, CFA offers researchers the ability to rigorously test hypothesized factor structures and compare model fit of competing factor structures (Jeon, 2015; Palmieri et al., 2007).

According to Armour and colleagues (2016), 112 CFA studies of *DSM*-correspondent PTSD measures were published between 1994 and 2016. A search of the literature conducted by the author on October 1, 2019 revealed that 33 more PTSD CFAs were published since the Armour et al. review.

Historical Overview of PTSD Factor Structure

Historically, CFA gained prominence following the release of an updated conceptualization of PTSD in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (*DSM-IV*; APA, 1994). *DSM-IV* PTSD was characterized as a 17-symptom construct which can occur following exposure to a traumatic event that involved actual or threatened death, physical injury, or threat to physical integrity of self or others wherein a person felt intense fear, helplessness, or horror. Symptoms were categorized into three clusters, including reexperiencing the trauma, avoidance of stimuli associated with the trauma and numbing of responsiveness, and increased arousal (APA, 1994).

Shortly following its release, the *DSM-IV's* implicit three-factor model was the subject of widespread criticism and intensive empirical scrutiny. Investigators proposed several alternative structures that differed in number of factors and symptom distribution, drawing on theoretical considerations or empirical CFA findings (Armour et al., 2016a). Item mappings for the various *DSM-IV* models are provided in Table 1. First, King, Leskin, King, and Weathers (1998) proposed a four-factor Emotional Numbing model, which separated effortful avoidance and emotional numbing into distinct clusters. The resulting avoidance cluster comprised the two effortful avoidance symptoms, whereas the resulting emotional numbing factor comprised traumatic amnesia, loss of interest, detachment, restricted affect, and sense of foreshortened future. The reexperiencing and hyperarousal clusters were unaltered, and thus the hyperarousal cluster continued to comprise sleep problems, irritability and anger, concentration problems, hypervigilance and exaggerated startle. This model was based on model fit from CFA and supported by conceptual judgments that avoidance and emotional numbing are distinct constructs (Asmundson, Gordon, Stapleton, & Taylor, 2004; Foa, Riggs, & Gershuny, 1995; King & King, 1994).

Next, Simms, Watson, and Doebbeling (2002) proposed the four-factor Dysphoria model on the theoretical basis that a fourth non-specific dysphoria factor was needed to capture symptoms that represent general distress. This model was founded on CFA findings that sleep problems, irritability and anger, and concentration problems clustered more closely with emotional numbing symptoms than did hypervigilance and exaggerated startle, and were strongly associated with external correlates of depression and anxiety (Simms et al., 2002; Taylor et al., 1998). Accordingly, the Dysphoria model created a dysphoria factor comprising the emotional numbing items, plus sleep problems, irritability and anger, and concentration problems. This left hypervigilance and exaggerated startle as the only symptoms in the hyperarousal factor.

Subsequently, the majority of *DSM-IV* structural validity studies focused on comparing the Emotional Numbing and Dysphoria models (Armour et al., 2016a). A meta-analysis of CFA findings from 40 studies revealed a small advantage of the Dysphoria model over the Emotional Numbing model across trauma types and PTSD measures (Yufik & Simms, 2010). In a systematic review of 111 factor analytic comparisons of the four-factor models, Armour and colleagues (2016a) found that 52 comparisons supported the Dysphoria model, 39 supported the Emotional Numbing model, and 19 observed no difference. In all of these CFA studies, both the Emotional Numbing and Dysphoria models consistently outperformed the *DSM-IV* model. However, they often resulted in suboptimal fit (Elhai et al., 2011). Also, the closeness in model fit between Emotional Numbing and Dysphoria models overall was problematic, given that they represent distinct conceptual and structural representations of PTSD (Elhai et al., 2011).

In a bid to resolve this dilemma, Elhai et al. (2011) proposed a five-factor Dysphoric Arousal model, which preserved both of the changes made in the Emotional Numbing and Dysphoria models. The Dysphoric Arousal model retained the *DSM-IV* reexperiencing factor, the Emotional Numbing avoidance factor, and the Dysphoria hyperarousal factor, which Elhai et al. (2011) renamed anxious arousal. However, instead of grouping sleep problems, irritability and anger, and concentration problems with either the hyperarousal factor or dysphoria factor, the Dysphoric Arousal model placed them in a new dysphoric arousal factor, leaving emotional numbing items and hyperarousal items in their own respective factors. The end result was a model with reexperiencing, avoidance, numbing, dysphoric arousal, and anxious arousal factors, which was founded on theory, CFA evidence, and treatment outcome research (Pietrzak et al., 2015; Shevlin, Hyland, Karatzias, Bisson, & Roberts, 2017). In Armour et al.'s (2016a) review of CFA comparisons, the five-factor Dysphoric Arousal model was superior to both four-factor models in 29 of 36 samples and provided worse fit to the Dysphoria model in only one sample. Further, the Dysphoric Arousal model outperformed *DSM-IV* in 20 of 22 samples.

Compared to these rival four-factor and five-factor models, the three-factor *DSM-IV* model was only found to be best-fitting in three of over 100 samples (Armour et al., 2016a). The *DSM-IV* model appeared superior to a unidimensional model of PTSD, and slightly, but not definitively, superior to two-factor models, fitting worse in11 samples and better in 19 samples. The consistency and extent of this evidence persuaded the *DSM-5* PTSD work group to replace the three-factor symptom structure with a four-factor, separating avoidance and numbing (Friedman, Resick, Bryant & Brewin, 2011; Friedman, 2013a; Friedman 2013b).

Other significant changes to the PTSD criteria for *DSM-5* included the addition of three symptoms and revision of several others. As a result of these changes, PTSD was reformulated as a 20-symptom construct with four factors (APA, 2013). Reexperiencing (REX) includes intrusive memories; nightmares; flashbacks; psychological reactivity to reminders; and physiological reactivity to reminders. Avoidance (AV) includes avoidance of trauma-related thoughts and feelings and avoidance of external trauma reminders. Negative alterations in cognition and mood (NACM) includes traumatic amnesia; negative cognitions about self, others, and the world; distorted blame; persistent negative emotions; loss of interest; detachment; and inability to experience positive emotions. Alterations in arousal and reactivity (AAR) includes irritable, angry, or aggressive behavior; reckless or self-destructive behavior; hypervigilance; exaggerated startle; concentration problems; and sleep problems. This structure is most similar to the four-factor Emotional Numbing model, which, as described above, was one of the leading models in the *DSM-IV* factor analytic literature. As noted by Weathers and colleagues (2014), these changes triggered questions regarding clinical implementation of the new criteria, revisions of *DSM*-correspondent assessment tools, correspondence between *DSM-IV* and *DSM-5* criteria, and construct validation of *DSM-5* PTSD.

Once the *DSM-5* PTSD criteria became available, the field embarked on a new round of CFA in attempt to evaluate the implicit four-factor *DSM-5* model. As expected, CFA results revealed support for the *DSM-5* model as a significant improvement over the previous *DSM-IV* three-factor model (Contractor et al., 2014; Elhai et al., 2011; Gentes et al., 2014; Liu et al., 2014; Miller, Wolf, & Keane, 2014).Despite that, CFAs also indicated that the four-factor *DSM-5* model was still inferior in fit compared to the five-factor Dysphoric Arousal model (Armour et al., 2016a; Elhai et al., 2011; Tsai et al., 2014). However, with the introduction of three new symptoms and substantial rewording of other symptoms for *DSM-5*, the Dysphoric Arousal model, which is based on *DSM-IV* criteria, no longer corresponded directly to the revised PTSD criteria. This propelled a new search for other, increasingly complex models that could accommodate the revised and expanded *DSM-5* criteria. These new models included the six-factor Anhedonia model (Liu et al., 2014), six-factor Externalizing Behavior model (Tsai et al., 2014), and seven-factor Hybrid model (Armour et al., 2015). Item mappings for these *DSM-5* models are provided in Table 2.

Both the six-factor Anhedonia model proposed by Liu and colleagues (2014) and six-factor Externalizing Behavior model proposed by Tsai and colleagues (2014) retain the separation of the *DSM-5* hyperarousal factor into dysphoric and anxious arousal clusters as in the *DSM-IV* Dysphoric Arousal model. The Anhedonia model divides NACM into negative affect and anhedonia factors based on previous construct validation of positive and negative affect as different concepts (Cuthbert & Kozak, 2013; Watson 2000, 2005; Watson, Gamez, & Simms, 2005). The Externalizing Behavior model includes an additional factor for *DSM-5* symptoms that represent self-initiated aggressive behaviors (e.g., irritable or angry behavior; reckless or self-destructive behavior). Tsai and colleagues (2014) argue that these observable, active behaviors are theoretically different from the remaining alterations in arousal and reactivity symptoms, which are subjective passive experiences. This assertion was supported by findings from a sample of 1,484 U.S. veterans from the National Health and Resilience in Veterans Study (NHRVS) that that externalizing behaviors and anhedonia had differential patterns of associations with external correlates such as depression, suicide, and hostility (Pietrzak et al., 2015). Lastly, the seven-factor Hybrid model developed by Armour and colleagues (2015) combines the changes made in both six-factor models, resulting in seven factors of PTSD symptoms: reexperiencing, avoidance, negative affect, anhedonia, externalizing behavior, anxious arousal, and dysphoric arousal.

Each of these competing six- and seven- factor models has substantial empirical support. First, in a sample of 1,196 Chinese earthquake survivors, all models demonstrated good fit, but the fit of the six-factor Anhedonia model of PTSD was excellent and significantly better than the four-factor *DSM-5* model, the four-factor Dysphoria *DSM-IV* and *DSM-5* models, and the Dysphoric Arousal *DSM-IV* and *DSM-5* models (Liu et al., 2014). CFAs on data from the nationally representative NHRVS indicated that the six-factor Externalizing Behavior model demonstrated superior fit to both four-factor *DSM-5* and five-factor Dysphoric Arousal models, regardless of gender of veteran or chronicity of PTSD (Tsai et al., 2014).

Second, in two independent samples of trauma-exposed veterans and undergraduate students, the four-factor *DSM-5* model demonstrated acceptable but significantly worse fit than competing six- and seven-factor models (Armour et al., 2015). Between the two six-factor models, Anhedonia outperformed Externalizing Behaviors in model fit. The best performing model was the seven-factor Hybrid model, which provided significantly better fit than all other models (Armour et al., 2015).

Finally, this trend of CFA findings is represented throughout the literature, including international samples (Murphy et al., 2018). In a more recent review of 23 *DSM-5* CFA papers utilizing 27 samples, Rasmussen, Verkuilen, Jayawickreme, Wu, and McCluskey (2018) found that the Hybrid model had superior in fit in 70% of samples. In sum, the empirical evidence to date clearly supports the seven-factor Hybrid model as having the greatest structural validity for *DSM-5* PTSD.

However, critics have raised several key concerns about the increasing complexity of PTSD CFA models, including a lack of theoretical justification, limited conceptual value, and lack of clinical utility for further dividing the *DSM-5* model into five, six, or seven factors. First, they note that the justification for splitting clusters is primarily empirical rather than conceptual. A conceptual justification would involve an extensive construct validation process, including a comprehensive investigation of differential associations with external correlates, rather than acceptance solely on the basis of superior model fit (Lee et al., 2019).

Second, they argue that that new factors proposed in the six- and seven-factor models are underrepresented by having only two items, resulting in the lack of content coverage (Lee et al., 2019; Palmieri, Marshall, & Schell, 2007; Rasmussen et al., 2018). Rasmussen and co-authors (2018) explain that two-item factors are either indicators of a legitimate latent factor, in which case more indicators (symptoms) are needed to properly identify it, or the two items are related in similarity or causality and are not indications of a latent phenomenon. Methodologists consider two-item factors underidentified and thus insufficient evidence that theoretically valid latent phenomena exist (Rasmussen et al., 2018). Practically, smaller numbers of indicators across factors (Armour et al., 2016a, 2015) also limit the type of statistical analyses that can be applied to evaluate these models (Kline, 2011; Marsh, Hau, & Wen, 2004).

Third, critics argue that the six- and seven-factor models have limited conceptual and incremental value. From this perspective, the same CFA findings used to justify six- and seven-factor models are taken as evidence that the four-factor *DSM-5* model typically achieves at least acceptable fit and could be retained as a supported and more parsimonious symptom structure (Weathers et al., 2014). They also reference known limitations of CFA---such as how increasing the number of factors will always increase the goodness of fit (Brown, 2015; Trninic, Jelaska, &

Stalec, 2013)---to question whether expanding the number of factors adds value beyond mathematical optimization (Rasmussen et al., 2018). As Rasmussen et al. (2018) explained, “with a fixed number of items and a push for more and smaller factors, the relative relevance of each factor to the PTSD construct likely decreases – no matter how coherent factors may be statistically.” (p. 241). The main concern here is locating the most parsimonious, empirically and conceptually supported construct, which would be the most useful for clinical diagnosis (Clark & Watson, 1995; Strauss & Smith, 2009).

This concern regarding the lack of additive conceptual value is further strengthened by examination of the observed correlations between factors. Rasmussen and colleagues' (2018) literature review aptly highlighted the magnitude of factor correlations across CFA studies of six- and seven-factor models, the majority of which were .70 or higher. Although general psychological distress would likely contribute to high associations between constructs in clinical psychopathology, factor correlations above .70 or .80 suggest that these factors substantially overlap, lack discriminant validity, and are not clinically distinct. For example, strong correlations among symptom clusters observed in several studies (Lee et al., 2019; Silverstein, Dieujuste, Kramer, Lee, & Weathers, 2017), such as strong associations between NACM items and dysphoric arousal items within the seven-factor Hybrid model, indicate that they may be better conceptualized as a single general distress construct, as represented in the *DSM-5* NACM cluster (Simms et al., 2002). Using Wald testing through latent variable modeling and external correlates from the Personality Assessment Inventory (PAI; Morey, 2007), Silverstein and colleagues (2017) found evidence supporting the differentiation between dysphoric arousal, anxious arousal, and externalizing behavior factors of the Hybrid model, but a lack of evidence for the differentiation between negative affect and anhedonia. Similarly, Lee and colleagues

(2019) found that the negative affect and anxious arousal factors in the Hybrid model did not yield differential associations with external correlates as would be expected if they were theoretically different constructs.

Fourth, critics raise the practical clinical issue that diagnostic algorithms based on a seven-factor solution might greatly restrict the possibility of meeting for PTSD diagnosis, thereby influencing treatment or service access and future research (Lee et al., 2019; Murphy et al., 2018; Shevlin et al., 2017). This concern is based on the assumption that the number of factors in a PTSD CFA model will translate directly into the same number of diagnostic criteria. However, this concern may not be well-founded. There is no necessary link between the number of factors identified in the construct of PTSD and the number of criteria required for a PTSD diagnosis. Although such a linkage occurred in moving from *DSM-IV* to *DSM-5* criteria---i.e., the separation of avoidance and numbing was accompanied by the addition of avoidance as a separate diagnostic criterion---this was an independent conceptual decision by the PTSD committee. A plausible alternative diagnostic rule, for example, would have been to separate avoidance and numbing into distinct clusters, but only require one or the other.

As a related example, multifactor models of major depression have been identified (see Huang & Chen, 2015 and Vanheule, Desmet, Groenvynck, Rosseel & Fontaine, 2008), yet depression symptoms are not clustered in *DSM*, and the diagnostic rule only requires that five of nine symptoms be present and that one of them be depressed mood or anhedonia. Nevertheless, a critique of increasingly complex PTSD models, based on the assumption that each factor will translate directly to a diagnostic criterion, is that increasing the number of factors will increases the number of criteria and thereby create an overly stringent diagnostic rule. For example, according to this line of reasoning, the Hybrid model would require that individuals have engaged in externalizing behavior to qualify for diagnosis, which trauma experts note is questionable (Armour et al., 2016a). Apropos of this concern, Shevlin and colleagues (2017) and Murphy and colleagues (2018) observed significant declines in prevalence rates of PTSD diagnosis when shifting from a *DSM-5* diagnostic system to models with more factors.

In summary, PTSD factor structure research is now at a conceptual impasse. Proponents of the seven-factor Hybrid model cite its repeated outperformance of the *DSM-5* model in CFAs across samples and measures. Critics argue that the Hybrid model has been over-fitted, creating superfluous factors with insufficient external validation, considerable overlap, and little incremental validity over the more parsimonious *DSM-5* model, which typically produces adequate fit in CFA. The debate is energized on one side by encouraging empirical evidence from CFA and supported on the other by philosophical arguments for theory, parsimony, and practical relevance.

Methodological Limitations of CFA

This conceptual gridlock is complicated further by a substantial limitation of CFA that is especially relevant to PTSD factor structure. Traditional CFA is also referred to as the independent clusters model CFA (ICM-CFA) because of its key statistical requirement that secondary factor loadings or "cross-loadings" are fixed to zero and not estimated (Assis Gomes, Almeida, & Núñez, 2017; Perry, Nicholls, Clough, & Crust, 2015; Prudon, 2015). As Assis Gomes and colleagues explained, "CFA demands that each latent variable of the measurement model loads exclusively on at least two items related to that latent variable […and…] must not load on any other latent variables" (pp. 397-8). This extreme restriction to zero cross-loadings imposed by the ICM makes CFA easier to interpret, but may often be unrealistic, if not frankly antithetical to contemporary conceptualizations of many complex constructs, including PTSD

(Assis Gomes et al., 2017; Marsh, 2007; Marsh, Morin, Parker, & Kaur, 2014). This is because observed indicators often predict significant variance in more than one latent construct (Rasmussen et al., 2018).

Assis Gomes and colleagues (2017) gave the examples of latent variables such as depression and anxiety, noting their significant overlap in predicted variance despite being widely accepted as different constructs. Likewise, participants' responses on any single PTSD item (e.g., anhedonia) might be most highly associated with its expected factor (e.g., NACM), but might also explain some nonzero quantity of covariance in another factor (e.g., avoidance) as well (Rasmussen et al., 2018). There is also theoretical support for symptoms being conceptually related to more than one construct in clinical psychopathology, due to influence of general distress (Rasmussen et al., 2018) and the frequent occurrence of psychiatric comorbidity between symptoms belonging even to separate diagnostic categories (Murray, Booth, Eisner, Obsuth, & Ribeaud, 2018). Marsh, Morin, Parker, and Kaur (2014) argued that nonzero cross-loadings are inherent in psychological constructs and will be true of most items. Also, small cross-loadings can be justified by item content (e.g., residual associations due to method effects) or measurement error (e.g., residual associations due to fallibility of indicators or interpretation by respondents; Asparouhov & Muthén, 2009; Marsh, Morin, Parker, & Kaur, 2014). In cases where cross-loadings are conceptually justified or, at the very least, expected to be nonzero, constraining cross-loadings to zero, as required by CFA, is a conceptually inappropriate approach to testing structural validity.

Second, the application of CFA when inappropriate may result in misspecifications for each nonzero cross-loading, which are errors between the prescribed model and estimated parameters (Perry, Nicholls, Clough, & Crust, 2015). Conducting an CFA that does not model cross-loadings is akin to suppressing any nonzero covariance, and this non-explained covariance is typically dispersed to factor correlations (Rasmussen et al., 2018). Misspecification of nonzero factor loadings can produce overestimated factor correlations (Prudon, 2015), distort the size of relations among factors (Marsh et al., 2009), and produce improper solution (Dillon, Kumar, & Mulani, 1987). Even small cross-loadings forced to zero can result in inflated CFA factor correlations that "detract from the discriminant validity of the factors and lead to biased estimates" (Marsh et al., 2014, p.88; also see Asparouhov & Muthén, 2009 and Marsh et al., 2009). One example is the chi-square (χ^2) test statistic, which is uniquely sensitive to sample size, resulting in poor tests of fit for models with small or trivial misspecifications in large samples (Asparouhov & Muthén, 2009; Perry et al., 2015). Model misspecifications also affect relative fit indices and non-centrality indices, although to different degrees, increasing the likelihood that even researchers who rely on the "selection-of-fit" indices approach to identify acceptability of fit (Brown, 2015; Perry et al., 2015) will make a Type II error and incorrectly reject a good model.

The more items or factors contained in a latent factor model, the greater the risk of substantial changes in model fit due to prevalent misspecifications (Perry et al., 2015). Item-level CFAs often do not support the predicted factor structures of complex constructs due to their overly restrictive assumptions detailed above (Marsh, Nagengast, & Morin, 2012 in Assis Gomes et al., 2017). Consequently, measures of complex constructs such as PTSD that are conceptually valid and well-performing in clinical practice and applied research, may achieve model fit well below accepted criteria (Marsh, 2007; Marsh et al., 2009, 2009, 2010). When this occurs, it can be difficult to know if the model is poor or if model misspecifications have resulted in Type II error, whereby a good model is rejected. This problem has been implied in the PTSD CFA literature, but not yet directly addressed. For example, there are repeated indications of barely adequate fit, high factor loadings, and inflated factor correlations in theoretically sound four-factor *DSM-5* CFA solutions, which suggests presence of the aforementioned model misspecifications (Gentes et al., 2014; Rasmussen et al., 2018).

Due to the ICM restriction, the use of CFA for complex multidimensional constructs such as PTSD may result in erroneous conclusions, leading to inappropriate conclusions regarding model acceptability. Despite warnings about the inapplicability of CFA to many psychological instruments and constructs, researchers continue to use CFA when inappropriate because it is widespread or they are unaware of other techniques that offer the same methodological advantages for latent constructs (Marsh et al., 2010). The use of ad hoc strategies to compensate for the use of CFA when inappropriate such as parceling or ex post facto modifications like ad hoc correlated uniqueness (Marsh et al., 2014) are also "dubious, counterproductive, misleading, or simply wrong" (Marsh et al., 2010, p.472), in that they can be applied atheoretically or cause bias in the results themselves. Even in their primer on the usefulness of CFA/SEM for the social sciences, Shanmugam and Marsh (2015) acknowledge the possibility of the rejection of well-justified theory and models on the basis of global fit statistics obtained via CFA or SEM. In these cases, they suggest checking for data errors and violations of SEM assumptions, fitting a just-identified model, which they concede is fruitless, or examining effect size, confidence intervals, and lower-order components to evaluate the model, leaving researchers with no practical solutions.

Researchers are faced with several choices when encountering misspecifications in the four-factor *DSM-5* PTSD model (Perry et al., 2015). Many, including supporters of the *DSM-5* model may conclude that the misspecifications are either unimportant or justified in the name of parsimony, and proceed with a four-factor conceptualization. In contrast, supporters of the seven-factor Hybrid model are more likely to conclude that the misspecifications are significantly relevant and proceed to reject or modify the model. To achieve acceptable fit, researchers may implement some or all suggested modification indices (MI) or choose a more complex conceptualization, perhaps as suggested by previous EFAs. Perry and colleagues (2015) suggest that all three avenues may be flawed. The first results in a theoretically appropriate model that seemingly lacks empirical support; the second leaves researchers with no model; the third data-driven approach to improving model fit may result in an empirically supported model lacking *a priori* theoretical support. This aptly parallels the current crossroads in PTSD factor structure research discussed above. It appears that an alternative method free from these limitations is needed to tackle unanswered PTSD factor structure questions.

Introduction to Exploratory Structural Equation Modeling (ESEM)

Exploratory structural equation modeling (ESEM; Asparouhov & Muthén, 2009) is a recently introduced analytic technique intended to address the existing limitations of CFA and EFA. Similar to CFA, ESEM is used to test fit of data to *a priori* predictions, such as invariance or a measurement model or relations between constructs (Howard, Gagne, Morin, & Forest, 2016 in Assis Gomes et al., 2017). It allows researchers to access the advantages of CFA such as *a priori* definition of latent and observed relationships, estimation of model fit, constructs adjusted for measurement error, bifactor models, complex error structures, group invariance analysis, and multiple-indicator multiple-cause (MIMC) models (Assis Gomes et al., 2017; Marsh et al., 2014; Morin & Maïano, 2011). Similar to EFA, ESEM does not require that items exclusively load onto only one specific latent variable (Asparouhov & Muthén, 2009; Assis Gomes et al., 2017; Perry et al., 2015). ESEM allows all factor loadings to be estimated given that the model is properly identified (Marsh et al., 2010; Perry et al., 2015), permitting all observed variables to load onto the latent factors. This relaxes the CFA restrictions, making the ESEM model more applicable to complex constructs (Marsh et al., 2014; Perry et al., 2015), and eliminates the possibility of secondary, often non-significant cross-loadings causing irrelevant model misspecification and potential rejection of a good model (Perry et al., 2015).

Identification. Researchers can identify all parameters in ESEM with maximum likelihood (ML) estimator, weighted least square estimators, or other robust estimators such as robust maximum likelihood (MLR) estimator (Marsh et al., 2014) according to what is appropriate for the data. Researchers can define factors as ESEM or CFA factors and group factors into blocks estimated by sets of indicators. Unlike in CFA, researchers can assign specific items to more than one set of factors, as indicated by theoretical predictions or empirical findings. Identification of ESEM mean structure parallels CFA in that latent factor means are constrained to zero while item intercepts are estimated by default. A single-factor ESEM model or ESEM model with only CFA factors specified is equivalent to a CFA model. When researchers are testing ESEM models with more than one factor ($m > 1$) with cross-loadings, they must first identify an unconstrained factor structure with m^2 constraints to be estimated. Then, the solution is rotated using an applicable rotation (Marsh et al., 2014).

Model Fit. Tests of ESEM models are completed the same ways as tests for regular SEM models (Asparouhov & Muthén, 2009), i.e., by using the standard χ^2 test of fit ($p \geq .05$), Standardized Root Mean Square Residual (SRMR < .08; Hu & Bentler, 1999), Bentler Comparative Fit Index (CFI ≥ .90; Bentler, 1990), Tucker-Lewis Index (TLI ≥ .90; Bentler, 1990), and Root Mean Square Error of Approximation (RMSEA; close fit if $p > .05$, poor fit if upper limit of 90% C. I. ≥ .10; (Browne & Cudeck, 1993; Kline, 2011). The basic CFA model is nested under the corresponding ESEM, as the CFA model can be obtained by placing restrictions on secondary factor loadings. Therefore, researchers can use conventional approaches to compare model fit. It is recommended that researchers rely on the sum of information from nested comparisons, model fit statistics, and parameter estimates for both models to make judgments about which model has superior fit (Asparouhov & Muthén, 2009; Brown, 2015; Browne & Cudeck, 1993; Marsh et al., 2014).

Rotational indeterminacy. It is important to note that for ESEM the chosen rotation for the solution will affect the pattern of cross-loadings and the size of the factor correlations, but not goodness of fit indices (Asparouhov & Muthén, 2009; Marsh et al., 2014). As Marsh et al. (2014) stated, rotation strategies in ESEM are mathematical maximizations of the structural relationships between variables and factors. Applied without forethought, their disregard for the appropriateness of factor correlation estimates can result in an incomprehensible factor solution. With geomin rotations, ESEM models are specified by entering all observed indicators by all latent factors, which is equivalent to conducting an EFA but with a prespecified number of latent factors. For a latent trait with an expected two-factor solution (F) and six observed indicators (v), an example specification might be:

```
ROTATION = GEOMIN(OBLIQUE, .5);
MODEL: F1-F2 by v1-v6 (*1);
```

Although early applications of ESEM used geomin rotations, the use of target rotation is now recommended (Marsh et al., 2014) for more *a priori* control of factor structure. Target rotation is a compromise between the data-driven mechanical rotation approach in EFA and the theory-driven specification of factors in CFA. It is most appropriate when a clear *a priori* structure has been defined. Target rotation can accommodate, but does not require, anchor items with zero nontarget factor loadings, which provide an *a priori* structure for specifying the model and facilitate interpretation. Target rotation has been shown to outperform geomin rotations when models have many latent factors, many nonzero elements, or high variable complexity (Asparouhov & Muthen, 2009).

For example, a latent trait may be expected to have two latent factors consisting of three observed indicators each, where the observed indicators are predicted to minimally cross-load. In this case, the example model specification includes cross-loadings that are not significantly different from zero as targets:

ROTATION = TARGET;

MODEL: F1 by v1 v2 v3 v4~0 v5~0 v6~0 (*1);

F2 by v4 v5 v6 v1~0 v2~0 v3~0 (*1);

Importantly, this does not fix the factor loading of v4-v6 on F1 to zero (as in ICM-CFA), but instead allows the factor loadings to be freely estimated while mathematically deriving a rotated solution that is closest to the specifications (in this case, cross-loadings close to zero).

If an *a priori* rationale exists regarding the magnitude of expected cross-loadings, syntax can be adjusted to specify a new value for target rotation:

MODEL: F1 by v1 v2 v3 v4~.5 v5~.5 v6~.5 (*1);

ESEM is most useful in contexts where the more parsimonious CFA model is not theoretically applicable or does not achieve good fit (Assis Gomes et al., 2017; Marsh et al., 2014, 2009, 2010). Perry and co-authors (2015) argue that this is especially the case for complex psychopathology constructs like personality or PTSD. ESEM provides flexibility to test similar hypotheses about structural validity when traditional CFA is not appropriate, but it is not intended to replace CFA (Asparouhov & Muthén, 2009; Marsh et al., 2014) or be used outside its own set of appropriate applications (Marsh et al., 2010). Booth and Hughes (2014) caution that ESEM may be highly useful to capture the theoretical and measurement complexity in personality and develop psychometric instruments, but it has not always proven more useful than CFA and should be used in conjunction rather than as a substitute. ESEM can also be extended for uses such as measurement invariance tests, latent mean comparisons, and differential item functioning evaluation (Marsh et al., 2014).

ESEM Applications

From the inception of ESEM until 2014, there were 87 empirical applications of ESEM in several areas of psychological science including personality; child, parent, and relationship behavior; academic experiences and attitudes; and values and beliefs (Marsh et al., 2014). A search of the literature conducted by the author October 1, 2019 revealed that 183 more studies were published since the Marsh et al. (2014) review using ESEM for scale development and construct validity analysis for psychological variables. In one example, Marsh and colleagues (2010) used ESEM to overcome a long-standing paradox in personality assessment, namely, that CFAs have failed to provide support for extensively validated personality constructs such as the Big Five factor structure supported by EFA. These results led Big Five researchers to question the appropriateness of CFA for the analysis of personality data and other to question the validity of highly supported and most widely used personality measures (Marsh et al., 2014). The authors noted that they applied ESEM to circumvent the restrictions of the CFA to which they attribute the negative CFA findings.

In a sample of 1,570 secondary school students who completed the NEO-Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992 in Marsh et al., 2010), Marsh et al. (2010) found that the traditional CFA model five-factor solution did not provide acceptable fit, whereas the corresponding ESEM solution fit much better, with goodness of fit statistics far superior to any in previous CFAs on the NEO-FFI. The authors found evidence of similar pattern of correlations between ESEM and CFA models, but slightly higher factor loadings and significantly larger factor correlations in the CFA model, providing evidence that the inappropriate use of CFA produced misspecifications that resulted in inflated factor correlations and loadings. They were also able to effectively utilize ESEM for examining NEO-FFI structure invariance over gender and time, which were methodological advantages previously thought to be achievable only with CFA. In other studies, even when both ESEM and CFA models demonstrated poor fit or there was no significant difference in fit between models, factor correlations were still substantially smaller in ESEM (see Marsh 2011a and 2011b in Marsh et al., 2014),

To date only one study has utilized ESEM in PTSD. Reichenheim et al. (2018) conducted ESEM and CFA models on the *DSM-IV*-based civilian version of the PTSD Checklist (PCL-C; Weathers, Litz, Huska, & Keane, 1994) in a sample of Brazilian post-partum women. Unsurprisingly, the *DSM-IV* three-factor CFA demonstrated borderline fit and several indications of model misspecification (e.g., factor correlations exceeding suggested values, item loadings reaching 1.0, and discriminant validity violations). Authors tested one- to five-factor ESEMs with geomin rotation, and with and without residual correlations specified. They arrived at a best-fitting two-factor solution – (1) reexperiencing and avoidance and (2) emotional numbing and hyperarousal – which provided excellent fit. Tests of nested model comparisons were not included, and the authors used ESEM in an exploratory rather than confirmatory manner. The limitations of this study, including a highly specific sample and use of a *DSM-IV* PTSD measure, render the findings less generalizable and informative for the current literature on the structure of *DSM-5* PTSD.

Given the current impasse in PTSD factor structure research, the present study is an examination of ESEM as an alternative and potential remedy to the limitations of CFA. The purpose of the present study was to compare ESEM and CFA models of *DSM-5* PTSD on model fit, factor loadings and cross-loadings, factor intercorrelations, and differential associations with external correlates. In order to keep the comparisons between ESEM and CFA simple and readily interpretable, the only existing CFA models examined were the four-factor *DSM-5* model, which serves as a baseline model, and the seven-factor Hybrid model, which has consistently been found to be the best-fitting model overall.

Hypotheses

The first four hypotheses involved the comparison between ESEM and CFA with respect to model fit, factor loadings, and factor intercorrelations.

Hypothesis 1: The four-factor *DSM-5* and seven-factor Hybrid CFAs would demonstrate average and good fit, respectively, consistent with the substantial CFA literature.

Hypothesis 2: ESEMs for both the *DSM-5* and Hybrid models would demonstrate better model fit than the corresponding CFA models.

Hypothesis 3: ESEM would reveal significant cross-loadings, thereby indicating the appropriateness of ESEM for examining the factor structure of PTSD.

Hypothesis 4: ESEM factor loadings and factor intercorrelations would be lower than in corresponding CFAs.

Hypotheses 5 and 6 involved convergent and discriminant validity of the additional factors found in the seven-factor Hybrid model, specifically their differential associations with external correlates. The related analyses used the PTSD Checklist for *DSM-5* (PCL-5; Weathers et al., 2013a) for PTSD latent variables and the Personality Assessment Inventory (PAI; Morey,

2007) for external correlates. Three PAI variables were used (described in detail below in the Data Analyses section), including PAI-Depression, PAI-Anxiety, and PAI-Externalizing Behavior. Given that the purpose of the present study was to explore the potential advantages of ESEM over CFA, these analyses were limited to only the new factors that emerge from splitting the four factors of the DSM-5 model into the seven factors of the Hybrid model. Specifically, the comparisons of interest were (a) PCL-Negative Affect versus PCL-Anhedonia (which emerged from splitting *DSM-5* NACM), and (b) the three pairwise comparisons between PCL-Externalizing Behavior, PCL-Anxious Arousal, and PCL-Dysphoric Arousal (which emerged from splitting *DSM-5* AAR). Latent factors are henceforth referred to by the measure acronym and scale or subscale name, e.g., PCL-Anhedonia and PAI-Depression. Specifically, between-cluster (Hypothesis 5) and within-cluster (Hypothesis 6) comparisons of associations with external correlates were conducted. For Hypothesis 5, the between-cluster approach compared the magnitude of the associations between two PTSD latent variables with the same external correlate. For Hypothesis 6, the within-cluster approach compared the magnitude of the associations between a single PTSD latent variable and two different external correlates.

Based on Silverstein et al.'s (2017) construct validity study of the seven-factor Hybrid model, between-cluster differential associations with external correlates were hypothesized for both ESEM and CFA as follows.

Hypothesis 5a: PCL-Anhedonia would have a higher association than PCL-Negative Affect with PAI-Depression.

Hypothesis 5b: PCL-Dysphoric Arousal would have a higher association than PCL-Externalizing Behavior with both PAI-Anxiety and PAI-Depression, but a lower association with PAI-Externalizing Behavior.

Hypothesis 5c: PCL-Dysphoric Arousal would have a higher association than PCL-Anxious Arousal with PAI-Depression.

Hypothesis 5d: PCL-Anxious Arousal would have a higher association than PCL-Externalizing Behavior with PAI-Anxiety, but a lower association with PAI-Externalizing Behavior.

Although no studies to date have examined within-cluster comparisons for the seven-factor Hybrid model, the following hypotheses were made, based on the content of the various PCL-5 subscales and PAI external correlates.

Hypothesis 6a: PCL-Negative Affect would be more strongly associated with both PAI-Depression and PAI-Anxiety than with PAI-Externalizing Behavior.

Hypothesis 6b: PCL-Anhedonia would be more strongly associated with PAI-Depression than with either PAI-Anxiety or PAI-Externalizing Behavior.

Hypothesis 6c: PCL-Externalizing Behavior would be more strongly associated with PAI-Externalizing Behavior than with either PAI-Depression or PAI-Anxiety.

Hypothesis 6d: PCL-Anxious Arousal would be more strongly associated with PAI-Anxiety than with PAI-Depression, and more strongly associated with both PAI-Anxiety and PAI-Depression than with PAI-Externalizing Behavior.

Hypothesis 6e: PCL-Dysphoric Arousal would be more strongly associated with PAI-Depression than with PAI-Anxiety, and more strongly associated with both PAI-Depression and PAI-Anxiety than with PAI-Externalizing Behavior.

Lastly, Hypothesis 7 was that the predicted patterns of correlations with external correlates would be more salient in ESEM than in CFA, as evidenced by larger effect sizes for the differences between correlations. In summary, the ESEM was expected to be retained as the superior model based on non-misspecification, goodness of fit, and construct validity.

Method

Participants were 2,685 undergraduate students enrolled in psychology courses at a southern public university who voluntarily participated for course credit. Participants opted in based on subjective report of having experienced a stressful life event. For the present study, participants were excluded from data analyses based on several indicators of low-effort responding including completion of less than 50% of required surveys (n=384), Personality Assessment Inventory (PAI; Morey, 2007) scores above recommend cutoffs for valid responding (n=335), and unusually fast response time of less than one second per question (n=18). Participants' responses on the Life Events Checklist for *DSM-5* (LEC-5; Weathers, Blake et al., 2013b) and written trauma narratives were reviewed independently by two graduate student raters for Criterion A status. Interrater agreement was substantial ($\kappa = 0.91$, $p < .01$), and rater disagreements were resolved through consultation with an expert in assessment of trauma exposure and PTSD. Of the remaining 1,948 participants, 541 were excluded because they did not identify having experienced any category of stressful life event on the LEC-5, and 268 were excluded because their narrative description of the index event did not meet Criterion A, leaving a final sample of 1,139 students.

Measures

In addition to a short demographic survey, participants completed the Life Events Checklist for *DSM-5* (LEC-5; Weathers et al., 2013b) which screens for 17 categories of traumatic experiences. In the present study, if participants responded affirmatively to one or more of the 17 event types, they were asked to indicate how the event happened (e.g., experienced directly, witnessed, learned about it, exposed during work duties) and identify and describe their worst event. Participants were then instructed to reference their worst event when completing the PTSD Checklist for *DSM-5* (PCL-5; Weathers et al., 2013a), an extensively validated self-report measure of PTSD. The PCL-5 has 20 items that correspond to *DSM-5* PTSD symptoms. Participants are instructed to rate how much they were bothered by each symptom over the past month, using a 5-point scale (0 = *Not at all* to 4 = *Extremely*). PCL-5 scores have consistently demonstrated strong internal consistency, test-retest reliability, and convergent and discriminant validity (Blevins et al., 2015; Bovin et al., 2016; Wortmann et al., 2016).

Participants also completed the Personality Assessment Inventory (PAI; Morey, 2007), a well-validated self-report measure of personality and psychopathology, consisting of 344 items rated on a 4-point scale (i.e., 0 = *false*, 1 = *slightly true*, 2 = *mainly true*, and 3 = *very true*). The PAI comprises 4 non-overlapping validity scales, 11 clinical scales, 5 treatment scales, and 2 interpersonal scales that produce highly valid and reliable scores (Boone, 1998; Boyle & Lennon, 1994; McDevitt-Murphy, Weathers, Flood, Eakin, & Benson, 2007; Morey, 2007). See Table 3 for scale-level descriptive statistics for all measures and Table 4 for item-level descriptive statistics.

Data Analyses

Missing data were handled using full information maximum likelihood (FIML; Enders, 2010; Schafer & Graham, 2002). All latent variable models were conducted on the PCL-5 with Mplus 7.0 using robust maximum likelihood (MLR; Brown, 2015; Chou & Bentler, 1995; Curran, West, & Finch, 1996), given that scores on most variables measuring psychopathology were expected to be non-normally distributed. Both four- and seven-factor ESEMs were specified using ESEM factors with both principal loadings freely estimated and cross-loadings estimated with targeted rotation at zero, according to best practices outlined by Asparouhov and Muthen (2009) and following several recent examples of ESEM on scales with well-established factor structure (see Guo et al., 2019; Tóth-Király, Bőthe, Rigó, & Gábor, 2017). See Figures 1-3 for illustrations of model structure and Appendix A for MPlus syntax for the seven-factor ESEM.

Model fit was assessed using multiple fit indices: χ^2 ($p \geq .05$), Standardized Root Mean Square Residual (SRMR < .08; Hu & Bentler, 1999), Bentler Comparative Fit Index (CFI ≥ .90; Bentler, 1990), Tucker-Lewis Index (TLI ≥ .90; Bentler, 1990), Root Mean Square Error of Approximation (RMSEA; close fit if $p > .05$, poor fit if upper limit of 90% C. I. ≥ .10; (M. W. Browne & Cudeck, 1993; Kline, 2011). Models were compared using nested model comparisons and robust χ^2 difference tests (Satorra & Bentler, 2001), the Akaike Information Criterion (AIC; Anderson & Burnham, 2000) and the Bayesian Information Criterion (BIC; Kass & Wasserman, 1995). AIC and BIC compare non-nested models, where lower values indicate higher likelihood of replicating in subsequent samples (Kline, 2011).

Following recommendations for assessing structural validity (Perry et al., 2015), factor loadings and cross-loadings were examined for conformity to theoretical models of PTSD. Specifically, the magnitude of factor loadings and proportion of items that loaded significantly onto intended and unintended factors were examined. Cross-loadings were examined for statistical significance and salience, following recommendations by Brown (2015) that loadings of at least .30-.40 are considered salient. Factor loadings of CFA models and the ESEM were compared to determine if there is evidence of model misspecification or statistically-driven inflated factor loadings.

In an effort to empirically test Marsh et al.'s (2014) proposition that inflated CFA factor correlations compromise convergent and discriminant validity, a comparison between the differential associations of ESEM and CFA factors with external correlates was conducted. Latent variables derived from PAI scale representing depression, anxiety, and externalizing behavior were included with the seven-factor ESEM and CFA models. The indicators from these external correlates were not allowed to cross-load onto any factors other than their primary factor. PAI-Depression was modeled using the three PAI Depression (DEP) subscales, Cognitive (DEP-C) Affective (DEP-A), and Physiological (DEP-P), as observed indicators. PAI-Anxiety was modeled using the three Anxiety (ANX) subscales, Cognitive (ANX-C), Affective (ANX-A), and Physiological (ANX-P). Finally, PAI-Externalizing Behavior was modeled using the Antisocial Features (ANT) parent scale, Aggression (AGG) parent scale, Alcohol Problems (ALC) scale, and Drug Problems (DRG) scale. Following Witte, Domino, and Weathers (2015) and Silverstein et al. (2017), the Wald test statistic was used to quantify the decrement in model fit that occurred when pairs of covariances were constrained to equality with one another. For example, one such comparison would evaluate whether model fit was significantly worse when the covariance between PCL-Anhedonia and PAI-Depression was constrained to equality with the covariance between PCL-Negative Affect and PAI-Depression. Because of the large number of comparisons, the Benjamini and Hochberg (1995) method was used to maintain the false discovery rate or type I error at .05. Cohen's q effect size of the difference in magnitude between each pair of correlations was calculated and interpreted according to guidelines provided by Cohen (1988): < .10 none, .10-.30 small effect, .30-.50 medium effect, > .50 large effect. Lastly, the pattern of differential associations of the latent PTSD factors with external correlates was compared between ESEM and CFA.

Results

The final sample of 1,139 students ranged in age from 18 to 40 years (M = 19.82 years; SD = 1.15) and was predominately female (75.70%, n = 862) and non-Hispanic (93.90% n = 1,069). In terms of race, 88.60% (n=1,009) identified as Caucasian, 6.40% (n=73) as African American, 2.90% (n = 33) as Asian, 1.4% (n = 16) multiracial, and less than 1% (n = 8) Native American or Pacific Islander. The breakdown of participant-identified index traumatic events was 28.45% (n = 324) transportation accident, 16.59% (n = 189) sexual assault, 14.84% (n =169) suicide, 9.66% (n = 110) natural disaster, 8.08% (n = 92) physical assault, and 6.06% (n = 69) serious accident, among others. Approximately 44.86% (n = 511) of participants directly experienced the index traumatic event, 18.96% (n = 216) witnessed it, and 36.17% (n = 412) learned about it happening to a close family member or friend. The prevalence of presumptive PTSD was 17.70% (n = 202). This was calculated by considering PCL-5 items rated as 2 = *Moderately* as symptoms endorsed, and then following the DSM-5 PTSD diagnostic rule (at least one REX symptom, one AV symptom, 2 NACM symptoms, and 2 AAR symptoms).

The pairwise proportion of data present in the dataset was 89-100%. As expected, PCL-5 items and PAI scales and subscales were negatively skewed. Internal consistency of all PCL-5 and PAI scales and subscales was good (α > .80) or acceptable (PCL-5 Externalizing behavior, α = .71; PAI Drug problems, α = .73). For descriptive statistics of scales and items used in analyses, see Tables 3 and 4 respectively.

Hypothesis 1: Comparative Model Fit for Four-Factor vs. Seven-Factor Models

Regarding Hypothesis 1, CFA results replicated the substantial literature in support of the superior fit of the seven-factor Hybrid model over the four-factor *DSM-5* model. The fit of the four-factor CFA was not adequate, while the fit of the seven-factor CFA was good according to most fit statistics and significantly better than the four-factor CFA according to Satorra-Bentler χ^2 difference testing.

Hypothesis 2: Comparative Model Fit for ESEM vs. CFA

Regarding Hypothesis 2, both four- and seven-factor ESEMs fit the data significantly better than their CFA counterparts. The fit of the four-factor ESEM was inadequate according to some and adequate according to other fit statistics. The seven-factor ESEM demonstrated good fit according to all fit statistics except χ^2 and had optimal fit compared to all other models, including the four-factor ESEM. It was also the most likely to replicate in subsequent samples, as evidenced by the lowest Akaike and Bayesian Information Criterion values of all four models. See Table 5 for model fit statistics for all models and Table 6 for Satorra-Bentler χ^2 difference test results.

Hypothesis 3: Cross-Loadings

In support of Hypothesis 3, there were many statistically significant, non-zero cross-loadings, but few salient cross-loadings in both four- and seven-factor ESEM models. For the best fitting model, the seven-factor ESEM, two salient cross-loadings above magnitude .30 were present: PCL-5 Item 15 (irritability or aggression; Externalizing Behavior) and Item 16 (reckless behavior; Externalizing Behavior) on PCL-Anhedonia. Two cross-loadings with magnitude at least half of the anchor loading on a particular factor were present: PCL-5 Item 2 (nightmares; Reexperiencing) and Item 12 (loss of interest; NACM) on PCL-Externalizing Behavior. These salient cross-loadings all appear to implicate the PCL-Externalizing Behavior factor in that two items (15&16) should have primary loadings on Externalizing Behavior but are also loading on Anehdonia, and two other items (2 & 12) should have primary loadings on other factors but are cross-loading onto Externalizing Behavior.

Additionally, there were 18 statistically significant, positive cross-loadings (magnitude .08 - .29) and 47 small, non-significant positive cross-loadings (magnitude .02 - .14) present in the seven-factor ESEM.

Hypothesis 4: Factor Loadings

Regarding Hypothesis 4, standardized factor loadings and intercorrelations were examined for all models to evaluate differences between ESEM and CFA models. First, the standardized factor loadings for the four- and seven-factor CFAs were all statistically significant and above .6, with the exception of the amnesia item (statistically significant and magnitude ~.40), which is characteristically lower in prevalence than other PTSD symptoms. Primary standardized factor loadings on PCL-Reexperiencing, PCL-NACM, and PCL-AAR for the four-factor ESEM were statistically significant but more variable in magnitude (range = .18 – 1.20). Notably, the standardized factor loading of Item 18 (startle) on PCL-AAR had a magnitude greater than 1.0, which is not necessarily problematic (Deegan, 1978; Joreskog, 1999). It may indicate a correlation very near 1.0 or multicollinearity in the data (Kenny, 1979). Given that PCL-AAR has six indicators and there were no observed problems such as nonconvergence or negative error variance (Kenny, 1979), it is unlikely that this indicates a problem with under-identification. In contrast with traditional CFA, this ESEM model allowed for the estimation of both primary and secondary loadings of items onto multiple factors. Contrary to expectations, the primary factor loadings of PCL-5 Item 6 (avoiding internal reminders) and Item 7 (avoiding external reminders) had non-significant, low magnitude primary factor loadings on PCL-Avoidance, and they cross-loaded more strongly on PCL-Reexperiencing and PCL-NACM.

For the seven-factor ESEM, all items loaded as expected onto primary factor loadings except PCL-5 Item 15 (irritability or aggression), for which there were stronger cross-loadings on PCL-Anhedonia and PCL-Anxious Arousal than the primary factor loading on PCL-Externalizing Behavior. Again, a standardized factor loading of magnitude slightly greater than 1.0 was observed for Item 10 (blame) on PCL-Negative Affect, but there was no other evidence of under-identification. See Table 7 and 8 for four-factor and seven-factor model standardized factor loadings, respectively.

In support of Hypothesis 4, the differences in magnitude of factor loadings between ESEM and CFA indicated likely inflation of CFA factor loadings. The number of standardized factor loadings above magnitude .70 was 17 of 20 for the four-factor CFA, and 8 of 20 for the four-factor ESEM. Likewise, there were 18 of 20 standardized factor loadings above magnitude .70 for the seven-factor CFA and 9 of 20 for the seven-factor ESEM. Also in support of Hypothesis 3, the seven-factor ESEM factor intercorrelations (M = .48) were substantially lower than the seven-factor CFA factors intercorrelations (M = .75). See Table 9 for standardized factor intercorrelations for seven-factor latent variable models.

Next, discriminant validity testing of factors in both ESEM and CFA models was conducted by first running the seven-factor ESEM and CFA models with additional CFA external correlates of PAI-Depression, PAI-Anxiety, and PAI-Externalizing Behavior. This created a ten-factor CFA model with all CFA factors (Figure 5) and a ten-factor ESEM model with seven ESEM PTSD factors and three CFA PAI factors (Figure 6) that could then be subjected to Wald testing by constraining associations between PTSD and PAI factors to equality and quantifying decrement in model fit. The ten-factor CFA had adequate fit according to some (SRMR = .04; RMSEA = .05, C.I .05-.05, p = .90; CFI = 0.95) but not all (χ^2 = 1296.32, df = 360, p < .01; TLI = 0.94) fit statistics. The ten-factor ESEM had good fit according to some (SRMR = .03; RMSEA = .04, C.I .04-.05, p = 1.0; TLI = 0.95; CFI = 0.97) but not all fit statistics (χ^2 = 885.28, *df* = 282, *p* < .01). For both the ESEM and CFA, standardized loadings for the observed indicators of the three external correlates were all statistically significant and above .60.

Hypothesis 5a-d: Between-Cluster Associations with External Correlates

To test Hypotheses 5a-d, the magnitudes of between-cluster correlations with external correlates were examined for both models, as shown in Tables 10 and 11. Also shown in Tables 10 and 11 are the significance tests for constraining pairs of correlations to equality (Wald W, *p* value) and effect size of the differences between the pairwise correlations (Cohen's q). The results of between-cluster comparisons for Hypotheses 5a-d were as follows. As expected, PCL-Anhedonia had a significantly stronger association than PCL-Negative Affect with PAI-Depression in both ESEM and CFA. As hypothesized, no other external correlates significantly differentiated PCL-Anhedonia and PCL-Negative Affect in the ESEM. However, inconsistent with hypotheses, PCL-Anhedonia had a stronger association than PCL-Negative Affect with both PAI-Externalizing Behavior and PAI-Anxiety in CFA.

Regarding Hypothesis 5b, as expected, PCL-Dysphoric Arousal had stronger associations than PCL-Externalizing Behavior with both PAI-Depression and PAI-Anxiety for both ESEM and CFA. Also, as hypothesized, PCL-Dysphoric Arousal had a weaker association than PCL-Externalizing Behavior with PAI-Externalizing Behavior in ESEM. The CFA results were inconsistent for this hypothesis, producing a nonsignificant between-cluster Wald W. Regarding Hypothesis 5c, as expected, PCL-Dysphoric Arousal had a stronger association than PCL-Anxious Arousal with PAI-Depression in both ESEM and CFA. Regarding Hypothesis 5d, as expected, PCL-Anxious Arousal had a stronger association than PCL-Externalizing Behavior with PAI-Anxiety in both models. Also, as expected, PCL-Anxious Arousal also had a weaker association than PCL-Externalizing Behavior with PAI-Externalizing Behavior in both ESEM and CFA.

Hypothesis 6a-e: Within-Cluster Associations with External Correlates

To test Hypotheses 6a-e, the magnitudes of within-cluster correlations with external correlates were examined for both models, as shown in Tables 10 and 11. The results of within-cluster comparisons for Hypotheses 6a-e were as follows. As expected, PCL-Negative Affect had significantly stronger associations with PAI-Depression and PAI-Anxiety than with PAI-Externalizing Behavior, in both ESEM and CFA. Regarding Hypothesis 6b, as expected, PCL-Anhedonia had a stronger association with PAI-Depression than with both PAI-Anxiety and PAI-Externalizing Behavior, but only in ESEM. In CFA, PCL-Anhedonia has a significantly stronger association with PAI-Depression than with PAI-Anxiety, but not with PAI-Externalizing behavior.

Regarding Hypothesis 6c, as expected, PCL-Externalizing Behavior had a significantly stronger association with PAI-Externalizing Behavior than with both PAI-Depression and PAI-Anxiety in ESEM. However, only the latter difference in association was significant for CFA. Also, contrary to the expectation of no difference, PCL-Externalizing Behavior had a stronger association with PAI-Depression than with PAI-Anxiety in ESEM. Regarding Hypothesis 6d, as expected, PCL-Anxious Arousal had a stronger association with PAI-Anxiety than with PAI-Depression, and a stronger association with PAI-Anxiety than with PAI-Externalizing Behavior, in both ESEM and CFA. Contrary to expectations, PCL-Anxious Arousal did not have a stronger association with PAI-Depression than with PAI-Externalizing Behavior in either ESEM or CFA. Regarding Hypothesis 6e, contrary to expectations, PCL-Dysphoric Arousal did not have a stronger association with PAI-Depression than with PAI-Anxiety or PAI-Externalizing Behavior in either CFA or ESEM, after using the Benjamini and Hochberg correction.

Hypothesis 7: Comparative Differential Associations for ESEM vs. CFA

Finally, to test Hypothesis 7, the consistency of findings with hypotheses 5a-d and 6-a-e, as well as the effect sizes of the differences in associations, were compared across ESEM and CFA. There were four instances in which a pairwise comparison was significant and consistent with the hypothesis for ESEM but not for CFA. There were three instances where hypothesized differences we not found for either ESEM or CFA. There was only one instance in which an unexpected difference in association was found for ESEM, which may indicate that the ESEM factors were better distinguishable from each other than originally hypothesized. Of the between- and within-cluster comparisons with significant Wald Ws for both ESEM and CFA, the effect size of hypothesized differences in pairwise associations was either similar (50%) or higher (50%) for ESEM compared to CFA. Thus, it appears that, relative to CFA, ESEM provided stronger evidence in support of the hypothesized patterns of differential associations with external correlates. In sum, the cumulative evidence in the present study indicated that the seven-factor ESEM should be retained as the superior model on the basis of non-misspecification, goodness of fit, and patterns of associations with external correlates.

Discussion

An extensive literature regarding CFA of PTSD has generated increasingly complex models of PTSD factor structure, producing much debate within the field. Critics have recently questioned the appropriateness of the independent clusters model (ICM) requirement of CFA that assumes zero cross-loadings and forces secondary factor variance to zero (Assis Gomes, Almeida, & Núñez, 2017; Perry et al., 2015; Prudon, 2015) for complex constructs like PTSD.

They argue that applying CFA to PTSD is unrealistic and has produced distortions in factor loadings, associations, and discrimination ability; improper solutions; and biased fit indices in CFAs of PTSD (Rasmussen et al., 2018). The primary concern is that researchers may have unintentionally rejected a good, theoretically sound model -- like the four-factor *DSM-5* -- based on misleading fit information. Despite the disadvantages of CFA for PTSD, no studies to date have empirically evaluated these concerns about misspecifications, nor evaluated a potential alternative approach. The present study aimed to compare the results of PTSD factor analysis from CFA and an alternative method, ESEM, which does not have the same restrictions as CFA but still allows for *a priori* specification. Overall, results indicated that ESEM is a valuable approach that addresses some of the main concerns regarding CFA.

First, the current findings provided quantitative evidence of misspecifications of CFA. The large number of non-zero cross-loadings found in the ESEMs provided compelling evidence of misspecification of CFA in that each of the many non-zero cross-loadings observed were forced to zero in the CFA models. As pointed out by Rasmussen and colleagues (2018), an inappropriate constraint of even small magnitude loadings to zero may cause the previously mentioned decrements in model fit and inflations of estimated parameters. As expected, both four- and seven-factor ESEMs outperformed corollary CFAs in model fit to the data. The optimal fit of ESEM for both models provided preliminary evidence that CFA misspecifications may be worsening its goodness of model fit. Next, the direct examination of parameters and intercorrelations between ESEM and CFA revealed that ESEM produced lower magnitude factor loadings for most items and substantially reduced factor intercorrelations. This was consistent with repeated observations of artificially inflated CFA factor loadings and factor intercorrelations in CFA PTSD studies. Within- and between-factor associations with external correlates demonstrated ESEM factors were more readily differentiated by external correlates than were CFA factors. This provides further evidence that CFA factors may be plagued by superfluous variance from misspecifications, resulting in poorer discrimination, whereas this is less of a concern with ESEM factors. Overall, these findings indicate that CFA may not be the optimal statistical technique for PTSD factor structure studies, and its application would likely result in misspecifications.

Next, the present results provided support for the use of ESEM as an alternative to CFA. Factor loadings for ESEM models generally conformed to theoretical expectations, more so for the seven-factor ESEM than the four-factor ESEM. The fit of both the four-factor *DSM-5* model and seven-factor Hybrid model was improved using ESEM. In the four-factor ESEM, PCL Items 6 (internal avoidance) and 7 (external avoidance) did not load as expected on PCL-Avoidance. Instead, they had significant and salient cross loadings on PCL-Reexperiencing and significant, but lower magnitude, cross-loadings on PCL-NACM. Interestingly, this finding is similar to the two-factor solution indicated by Reichenheim et al.'s (2018) exploratory ESEM in that it suggests the collapse of avoidance items into the reexperiencing cluster. However, this finding was not substantiated by the seven-factor ESEM.

In the seven-factor ESEM, there were very few salient cross-loadings, indicating good overall conformity to the seven-factor conceptualization. Interestingly, the few salient cross-loadings present were for items that cross-loaded onto PCL-Externalizing Behavior or PCL-Externalizing Behavior indicators that cross-loaded onto other latent factors. This suggests that newly added items that constitute the PCL-Externalizing Behavior factor are not performing as cleanly as expected, and that further explication of the underlying construct is needed. The large number of non-zero cross-loadings present in ESEMs was consistent with theoretical expectations of at least some interrelatedness of PTSD symptoms and effects of general distress. Taken together the findings demonstrate ESEM may be more statistically and conceptually appropriate for modeling PTSD.

The present findings replicate the superiority of the seven-factor Hybrid model using a more conceptually and statistically appropriate technique than CFA – ESEM. The seven-factor ESEM performed significantly better than all other models in goodness of fit with the data. Lowered factor intercorrelations in ESEM compared to CFA somewhat bolster the construct validity of the seven factors in the Hybrid model against criticisms that they are too highly overlapping to be considered unique constructs. Additionally, the small number of salient cross-loadings present in the seven-factor ESEM provide some justification of the split from the *DSM-5* NACM and AAR factors into Negative Affect, Anhedonia, Dysphoric Arousal, and Anxious Arousal in the seven-factor model. Results from the present study substantiate the continued use of the seven-factor Hybrid model in factor structure of PTSD research, while assuaging concerns that this conclusion had been previously dependent on misspecification CFA results.

Despite the aforementioned strong performance of the seven-factor Hybrid model in ESEM, it is important to note that the findings of the present study suggested mixed evidence of construct validity. Although the present study used parent PAI scales instead of specific subscales, the differential associations of the seven factors with external correlates essentially replicated Silverstein et al.'s (2017) mixed findings. First, PCL-Negative Affect and PCL-Anhedonia demonstrated at least modest evidence of both within- and between-cluster differential patterns of associations with external correlates. This provides at least limited support of the distinctiveness for these new clusters formed by splitting *DSM-5* NACM.

In regards to the factors formed by separating *DSM-5* AAR, there were very clear differences in differential associations that substantiate the distinction of PCL-Externalizing Behaviors as a separate factor from other factors subsumed within the *DSM-5* PCL-AAR factor (i.e. PCL-Anxious Arousal and PCL-Dysphoric Arousal). This finding is consistent with the extensive literature establishing that internalizing and externalizing psychopathology are distinct (see Van Voorhees et al., 2014 for in-depth review) and questioning whether distinguishing externalizing and internalizing types is useful for PTSD (Elhai et al., 2011; Forbes, Elhai, Miller & Creamer, 2010; Tsai et al., 2014). The exception to this was the presence of four cross-loadings, all implicating the PCL-Externalizing Behaviors factor, which, as mentioned above, suggest the need for further evaluation of the construct and item performance. PCL-Dysphoric Arousal and PCL-Anxious Arousal exhibited consistent differentiation from PCL-Externalizing Behavior, but less so with each other. PCL-Anxious Arousal demonstrated weak evidence, and PCL-Dysphoric Arousal demonstrated very limited evidence of differential associations with external correlates. Further, the effect sizes of the differences between associations of PTSD factors with external correlates were relatively modest, most being small to medium according to Cohen (1988).

In the current study, superior model fit, few salient cross-loadings, and lower factor intercorrelations supported the seven-factor model, but only some of the new splits of the seven-factor model were supported by differential associations with external correlates. Similarly to Silverstein et al. (2018), PCL-Negative Affect and PCL-Anhedonia appear to be performing well as unique factors. More research is needed to expand upon the present results, but so far there is weaker evidence that Negative Affect and Anhedonia and Dysphoric and Anxious Arousal are well-differentiated from each other. Combined with evidence from Silverstein et al.'s (2017) construct validation of the seven-factor model, this raises questions about whether certain factor splits from the four-factor *DSM-5* to seven-factor Hybrid model are justified.

In total, the present findings provide strong evidence that substantiates concerns about atheoretical application and resulting misspecifications of CFA in PTSD. It also provides preliminary example and evidence for a potential statistical alternative in ESEM, which appears to produce superior model fit and remediate most misspecification concerns of CFA. That said, although the current study demonstrates that the more flexible modeling approach of ESEM better reflects the latent structure of PTSD, our results also alleviate concerns that the restrictive CFA model and its misspecifications may have contributed to drive for increasingly complex models of PTSD. Specifically, this study provided evidence of the superior fit of the seven-factor Hybrid model even when using an improved statistical approach.

In general, ESEM offers flexibility in defining a measurement model on a continuum from purely exploratory (EFA) to purely confirmatory (CFA). That is, researchers can opt to specify just the number of a prior latent factors with no conditions for loadings, any number and variation of primary loadings with targeted or specific cross-loading predictions, or primary loadings with fully constrained cross-loadings (equivalent to CFA). This is particularly advantageous for PTSD, as a construct with a well-established theorized four-factor structure but ill-fitting four-factor CFAs. Similar to the affordances of CFA, researchers can use ESEM for nested models, measurement and structural invariance, test-retest or longitudinal analyses, and scale development, all of which are crucial for ongoing construct validation of PTSD and DSM-correspondent measures. Adopting the use of ESEM for PTSD will also provide additional data regarding cross-loadings, for which researchers can begin to examine and make predictions. Being able to evaluate data about item performance and inter-item and inter-factor relationships in many samples and with many measures would be especially helpful for PTSD, a construct for which the conceptualization (broad vs. narrow) and number of symptoms is still being debated.

Lastly, with less risk of producing misspecifications like CFA studies, PTSD researchers can be more assured that their conclusions from latent variable modeling are not solely an artifact of statistical problems, but reflect substantive information about the performance of PTSD. Using ESEM instead of CFA will produce latent factors more truly representative of the PTSD construct, that in turn are more likely to perform as predicted in subsequent analyses. This will allow scholars to make more precise judgments about the goodness of model fit, comparisons of *DSM-5* and other empirically-derived six- and seven-factor models, and relations between variables of interest, as well as compare across the field to arrive at consensus. Though existing research on ESEM in latent variable analysis of psychopathology is limited, the current findings support the use of ESEM for PTSD construct validity studies. Based on the results, ESEM may be particularly well-suited for PTSD latent variable modeling moving forward, because it can directly address impediments in the CFA literature.

The present study has some notable limitations. First, the study relied on cross-sectional questionnaire responses as observed indicators of PTSD symptomatology. This may have contributed to shared method variance and thus higher correlations between PTSD factors and external correlates. A different pattern of results may have been obtained with clinician-rated interviews. This limitation is mitigated somewhat by recent evidence that method effects are relatively small in structural studies of PTSD (Lee et al., 2019). Further, to improve the quality of self-rated data, participants were excluded for inattentive or low-effort responding. Future studies of ESEM for PTSD should evaluate clinician ratings from structured PTSD interviews measures such as the Clinician-Administered PTSD Scale for *DSM-5* (CAPS-5; Weathers et al., 2013c).

Second, participants were an undergraduate convenience sample of primarily white, female university students, which may not generalize to a general sample of trauma-exposed adults. Most notably, the influence of the undergraduate sample was observed in the restriction of range on antisocial traits and several already low base-rate PTSD symptoms like physical aggression, which are likely to be more prevalent in treatment-seeking samples. Future studies should replicate and extend the use of ESEM to a wide range of trauma-exposed samples.

To our knowledge, the confirmatory approach of ESEM had not previously been used in PTSD. It is important to note that authors are still expanding upon the correct and best uses of ESEM (Marsh, Guo, Dicke, Parker, & Craven, 2019; Sellbom & Tellegen, 2019). Future research with ESEM on other PTSD measures and in other samples will determine if the results are replicated. However, the consistent finding of inflated factor intercorrelations across PTSD CFA studies implies that CFAs have been routinely misspecified for PTSD and ESEM can be predicted to fare better. In conclusion, the results of the present study demonstrated researchers need not be wedded to CFA for its practicality, especially once they are aware of a sounder alternative that is just as feasible. Not only does ESEM offer the methodological innovations of CFA, but it also bestows several advantages for PTSD. Given the current gridlock in the field of PTSD structural validity, the continued and creative use of ESEM would contribute significantly to the ongoing conceptualization of PTSD.

References

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

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Washington, DC: American Psychiatric Association.

Anderson, D. R., & Burnham, K. P. (2000). Null hypothesis testing: Problems, prevalence, and an alternative. *Journal of Wildlife Management*, *64*(4), 912. doi:10.2307/3803199

Armour, C., Contractor, A., Shea, T., Elhai, J. D., & Pietrzak, R. H. (2016b). Factor structure of the PTSD checklist for DSM-5: Relationships among symptom clusters, anger, and impulsivity. *Journal of Nervous and Mental Disease*, *204*(2), 108–115. doi:10.1097/NMD.0000000000000430

Armour, C.,Műllerová, J., & Elhai, J. D. (2016a). A systematic literature review of PTSD's latent structure in the Diagnostic and Statistical Manual of Mental Disorders: DSM-IV to DSM-5. *Clinical Psychology Review*, *44*, 60–74. doi:10.1016/j.cpr.2015.12.003

Armour, C., Tsai, J., Durham, T. A., Charak, R., Biehn, T. L., Elhai, J. D., & Pietrzak, R. H. (2015). Dimensional structure of DSM-5 posttraumatic stress symptoms: Support for a hybrid Anhedonia and Externalizing Behaviors model. *Journal of Psychiatric Research*, *61*, 106–113. doi:10.1016/j.jpsychires.2014.10.012

Asmundson, G. J. G., Stapleton, J. A., & Taylor, S. (2004). Are avoidance and numbing distinct PTSD symptom clusters? *Journal of Traumatic Stress*, *17*, 467-475. doi:10.1007/s10960-004-5795-7

Asparouhov, T., & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438. doi:10.1080/10705510903008204

Assis Gomes, C. M., Almeida, L. S., & Núñez, J. C. (2017). Rationale and applicability of Exploratory Structural Equation Modeling (ESEM) in psychoeducational contexts. *Psicothema*, (29.3), 396–401. doi:10.7334/psicothema2016.369

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*, 289-300. doi:0035-9246/95/57289

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. doi:10.1037/0033-2909.107.2.238

Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., & Domino, J. L. (2015). The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress*, *28*(6), 489–498. doi:10.1002/jts.22059

Boone, D. (1998). Internal consistency reliability of the Personality Assessment Inventory with psychiatric inpatients. *Journal of Clinical Psychology*, *54*, 839-43. doi:10.1002/(sici)1097-4679(199810)54:6<839::aid-jclp12>3.0.co;2-m

Booth, T. & Hughes, D. J. (2014). Exploratory structural equation modeling of personality data. *Assessment*, *21*, 260-271. Doi:10.1177/1073191114528029

Bovin, M. J., Marx, B. P., Weathers, F. W., Gallagher, M. W., Rodriguez, P., Schnurr, P. P., & Keane, T. M. (2016). Psychometric properties of the PTSD Checklist for Diagnostic and

Statistical Manual of Mental Disorders–Fifth Edition (PCL-5) in veterans. *Psychological Assessment*, *28*(11), 1379–1391. doi:10.1037/pas0000254

Boyle, G. J. & Lennon, T. J. (1994). Examination of the reliability and validity of the Personality Assessment Inventory. *Journal of Psychopathology and Behavioral Assessment, 16,* 173-87. doi:10.1007/BF02229206

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research: Second Edition*. New York, NY: The Guilford Press.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Chou, C. P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37–55). Thousand Oaks, CA: Sage.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309–319. doi:10.1037/1040-3590.7.3.309

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd Edition). Hillsdale, NJ: Erlbaum.

Contractor, A. A., Durham, T. A., Brennan, J. A., Armour, C., Wutrick, H. R., … & Elhai, J. D. (2014). DSM-5 PTSD's symptom dimensions and relations with major depression's symptom dimensions in a primary care sample. *Psychiatry Research*, *215*, 146–153. doi:10.1016/j.psychres.2013.10.015

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*(1), 16–29. doi:10.1037/1082-989X.1.1.16

Cuthbert, B. N., & Kozak, M. J. (2013). Constructing constructs for psychopathology: The NIMH research domain criteria. *Journal of Abnormal Psychology*, *122*, 928–937. doi:10.1037/a0034028

Dillon, W. R., Kumar, A., & Mulani, N. (1987). Offending estimates in covariance structure analysis: Comments on the causes of and solutions to Heywood cases. *Psychological Bulletin*, *101*, 126-135. doi:10.1037/0033-2909.101.1.126

Elhai, J. D., Biehn, T. L., Armour, C., Klopper, J. J., Frueh, B. C., & Palmieri, P. A. (2011). Evidence for a unique PTSD construct represented by PTSD's D1–D3 symptoms. *Journal of Anxiety Disorders*, *25*(3), 340–345. doi:10.1016/j.janxdis.2010.10.007

Elhai, J. D., & Palmieri, P. A. (2011). The factor structure of posttraumatic stress disorder: a literature update, critique of methodology, and agenda for future research. *Journal of Anxiety Disorders*, *25*(6), 849–854. doi:10.1016/j.janxdis.2011.04.007

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Foa, E. B., Riggs, D. S., & Gershuny, B. S. (1995). Arousal, numbing, and intrusion: Symptom structure of PTSD following assault. *The American Journal of Psychiatry*, *152*, 116-120. doi:10.1176/ajp.152.1.116

Forbes, D., Elhai, J. D., Miller, M. W., & Creamer, M. (2010). Internalizing and externalizing classes in posttraumatic stress disorder: A latent class analysis. *Journal of Traumatic Stress*, *23*, 340-49. doi:10.1002/jts.20526

Friedman, M. J. (2013a). Finalizing PTSD in DSM-5: Getting here from there and where to go next. *Journal of Traumatic Stress*, *26*(5), 548–556. doi:10.1002/jts.21840

Friedman, M. J. (2013b). PTSD in the DSM-5: Reply to Brewin (2013), Kilpatrick (2013), and Maercker and Perkonigg (2013). *Journal of Traumatic Stress*, *26*(5), 567–569. doi:10.1002/jts.21847

Friedman, M. J., Resick, P. A., Bryant, R. A., & Brewin, C. R. (2011). Considering PTSD for DSM-5. *Depression and Anxiety*, *28*(9), 750–769. doi:10.1002/da.20767

Gentes, E. L., Dennis, P. A., Kimbrel, N. A., Rissling, M. B., Beckham, J. C., & Calhoun, P. S. (2014). DSM-5 posttraumatic stress disorder: Factor structure and rates of diagnosis. *Journal of Psychiatric Research*, *59*, 60–67. doi:10.1016/j.jpsychires.2014.08.014

Guo, J., Marsh, H. W., Parker, P. D., Dicke, T., Lüdtke, O., & Diallo, T. M. O. (2019). A systematic evaluation and comparison between Exploratory Structural Equation Modeling and Bayesian Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, Advance Online Publication. doi:10.1080/10705511.2018.1554999

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. doi:10.1080/10705519909540118

Huang, C., & Chen, J. H. (2015). Meta-analysis of the factor structures of the Beck Depression Inventory-II. *Assessment*, *22*, 459–472. doi:10.1177/1073191114548873

Jeon, J. (2015). The strengths and limitations of the statistical modeling of complex social phenomenon: Focusing on SEM, path analysis, or multiple regression models.

International Journal of Social, Behavioral, Educational, Economic, Business, and Industrial Engineering, *9*, 1634-42. doi:scholar.waset.org/1999.10/10001434

Kass, R. E., & Wasserman, L. (1995). A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. *Journal of the American Statistical Association, 90*(431), 928–934. doi:10.2307/2291327

Kilpatrick, D. G. (2013). The DSM-5 got PTSD right: Comment on Friedman (2013). *Journal of Traumatic Stress*, *26*(5), 563–566. doi:10.1002/jts.21844

King, L. A. & King, D. W. (1994). Latent structure of the Mississippi Scale for Combat-Related Post-Traumatic Stress Disorder: Exploratory and higher order confirmatory factor analyses. *Assessment*, *1*, 275-291. doi:10.1080/0969594940010301

King, D. W., Leskin, G. A., King, L. A., & Weathers, F. W. (1998). Confirmatory factor analysis of the clinician-administered PTSD Scale: Evidence for the dimensionality of posttraumatic stress disorder. *Psychological Assessment*, *10*(2), 90–96. doi:10.1037/1040-3590.10.2.90

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Lee, D. J., Bovin, M. J., Weathers, F. W., Palmieri, P. A., Schnurr, P. P., Sloan, D. M., … Marx, B. P. (2019). Latent factor structure of DSM–5 posttraumatic stress disorder: Evaluation of method variance and construct validity of novel symptom clusters. *Psychological Assessment*. doi:10.1037/pas0000642

Liu, P., Wang, L., Cao, C., Wang, R., Zhang, J., Zhang, B., … Elhai, J. D. (2014). The underlying dimensions of DSM-5 posttraumatic stress disorder symptoms in an epidemiological sample of Chinese earthquake survivors. *Journal of Anxiety Disorders*, *28*(4), 345–351. doi:10.1016/j.janxdis.2014.03.008

Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport and exercise psychology. In *Handbook of sport psychology, 3rd ed* (pp. 774–798). Hoboken, NJ, US: John Wiley & Sons Inc.

Marsh, H. W., Guo, J., Dicke, T., Parker, P., & Craven, R. G. (2019). Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM) and SET-ESEM: Optimal balance between goodness of fit and parsimony. *Multivariate Behavioral Research*. Advance online publication. doi:10.1080/00273171.2019.1602503

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341. doi:10.1207/s15328007sem1103_2

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur, G. (2014). Exploratory Structural Equation Modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*(1), 85–110. doi:10.1146/annurev-clinpsy-032813-153700

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory Structural Equation Modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 439–476. doi:10.1080/10705510903008220

Marsh, H. W., Muthén, B., Morin, A. J. S., Lüdtke, O., Asparouhov, T., Trautwein, U., & Nagengast, B. (2010). A new look at the Big Five factor structure through Exploratory Structural Equation Modeling. *Psychological Assessment, 22*(3), 471–491. doi:10.1037/a0019227

McDevitt-Murphy, M. E., Weathers, F. W., Flood, A. M., Eakin, D. E., & Benson, T. A. (2007). The utility of the PAI and the MMPI-2 for discriminating PTSD, depression, and social phobia in trauma-exposed college students. *Assessment*, *14*, 181-195. doi:10.1177/1073191106295914

Miller, M. W., Wolf, E. J., & Keane, T. M. (2014). Posttraumatic Stress Disorder in DSM-5: New criteria and controversies. *Clinical Psychology: Science and Practice*, *21*(3), 208–220. doi:10.1111/cpsp.12070

Morey, L. C. (2007). *Personality Assessment Inventory (PAI): Professional manual (2nd ed.)*. Lutz, FL: Psychological Assessment Resources, Inc.

Morin, A. J. S., & Maïano, C. (2011). Cross-validation of the short form of the physical self-inventory (PSI-S) using Exploratory Structural Equation Modeling (ESEM). *Psychology of Sport and Exercise*, *12*(5), 540–554. doi:10.1016/j.psychsport.2011.04.003

Murphy, S., Hansen, M., Elklit, A., Yong Chen, Y., Raudzah Ghazali, S., & Shevlin, M. (2018). Alternative models of DSM-5 PTSD: Examining diagnostic implications. *Psychiatry Research*, *262*, 378–383. doi:10.1016/j.psychres.2017.09.011

Murray, A. L., Booth, T., Eisner, M., Obsuth, I., & Ribeaud, D. (2018). Quantifying the strength of general factors in psychopathology: A comparison of CFA with maximum likelihood estimation, BSEM, and ESEM/EFA Bifactor Approaches. *Journal of Personality Assessment*, 1–12. doi:10.1080/00223891.2018.1468338

Palmieri, P. A., Marshall, G. N. & Schell, T. L. (2007). Confirmatory factor analysis of posttraumatic stress symptoms in Cambodian refugees. *Journal of Traumatic Stress*, *20*, 2017-16. doi:10.1002/jts.20196

Palmieri, P. A., Weathers, F. W., Difede, J., & King, D. W. (2007). 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*, 329-41. doi:10.1037/0021-843X.116.2.329

Perry, J. L., Nicholls, A. R., Clough, P. J., & Crust, L. (2015). Assessing model fit: Caveats and recommendations for confirmatory factor analysis and exploratory structural equation modeling. *Measurement in Physical Education and Exercise Science*, *19*(1), 12–21. doi:10.1080/1091367X.2014.952370

Pietrzak, R. H., Tsai, J., Armour, C., Mota, N., Harpaz-Rotem, I., & Southwick, S. M. (2015). Functional significance of a novel 7-factor model of DSM-5 PTSD symptoms: Results from the National Health and Resilience in Veterans Study. *Journal of Affective Disorders*, *174*, 522–526. doi:10.1016/j.jad.2014.12.007

Prudon, P. (2015). Confirmatory Factor Analysis as a tool in research using questionnaires: A critique. *Comprehensive Psychology*, *4*, 10. doi:10.2466/03.CP.4.10

Rasmussen, A., Verkuilen, J., Jayawickreme, N., Wu, Z., & McCluskey, S. T. (2018). When did posttraumatic stress disorder get so many factors? Confirmatory factor models since DSM–5. *Clinical Psychological Science*, *7*, 234-248. doi:10.1177/2167702618809370

Reichenheim, M. E., Oliveira, A. G., Moraes, C. L., Coutinho, E. S., Figueira, I., & Lobato, G. (2018). Reappraising the dimensional structure of the PTSD Checklist: Lessons from the

DSM-IV-based PCL-C. *Revista Brasileira de Psiquiatria*, *40*(2), 154–162. doi:10.1590/1516-4446-2017-2239

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507–514.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177. doi:10.1037/1082-989X.7.2.147

Sellbom, M. & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment*. Advance online publication. doi:10.1037/pas0000623

Shanmugam, V. & Marsh, J. E. (2015). Application of structural equation modeling to the social sciences: A brief guide for researchers. *Mesure et évaluation en education*, *37*, 99-123. doi:10.7202/1036329ar

Shevlin, M., Hyland, P., Karatzias, T., Bisson, J. I., & Roberts, N. P. (2017). Examining the disconnect between psychometric models and clinical reality of posttraumatic stress disorder. *Journal of Anxiety Disorders*, *47*, 54–59. doi:10.1016/j.janxdis.2017.02.006

Silverstein, M. W., Dieujuste, N., Kramer, L. B., Lee, D. J., & Weathers, F. W. (2017). Construct validation of the hybrid model of posttraumatic stress disorder: Distinctiveness of the new symptom clusters. *Journal of Anxiety Disorders*, *54*, 17–23. doi:10.1016/j.janxdis.2017.12.003

Simms, L. J., Watson, D., & Doebbeling, B. N. (2002). Confirmatory factor analyses of posttraumatic stress symptoms in deployed and nondeployed veterans of the Gulf War. *Journal of Abnormal Psychology*, *111*, 637–647. doi:10.1037/0021-843X.111.4.637

Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology*, *5*, 1–25. doi:10.1146/annurev.clinpsy.032408.153639

Taylor, S., Kuch, K., Koch, W. J., Crocket, D. J., & Passey, G. (1998). The structure of posttraumatic stress symptoms. *Journal of Abnormal Psychology*, *107*, 154-160. doi:10.1037/0021-843X.107.1.154

Tóth-Király I., Bőthe, B., Rigó, A., & Gábor, O. (2017). An illustration of the Exploratory Structural Equation Modeling (ESEM) framework on the Passion Scale. *Frontiers in Psychology*, *8*, 1-15. doi:10.3389/fpsyg.2017.01968.

Trninic, V., Jelaska, I., & Stalec, J. (2013). Appropriateness and limitations of factor analysis methods utilised in psychology and kinesiology: Part II. *Fizicka Kultura*, *67*, 5–17. doi:10.5937/fizkul1202077T

Tsai, J., Harpaz-Rotem, I., Armour, C., Southwick, S. M., Krystal, J. H., & Pietrzak, R. H. (2014). Dimensional structure of DSM-5 posttraumatic stress disorder symptoms: Results From the National Health and Resilience in Veterans Study. *The Journal of Clinical Psychiatry*, *76*, 546–553. doi:10.4088/JCP.14m09091

Vanheule, S., Desmet, M., Groenvynck, H., Rosseel, Y., & Fontaine, J. (2008). The factor structure of the Beck Depression Inventory-II: An evaluation. *Assessment*, *15*, 177–187. doi:10.1177/ 1073191107311261

Van Voorhees, E. E., Dennis, P. A., Elbogen, E. B., Clancy, C. P., Hertzberg, M. A. … & Calhoun, P. S. (2014). Personality Assessment Inventory internalizing and externalizing structure in veterans with posttraumatic stress disorder: Associations with aggression. *Aggressive Behavior*, *40*, 582-592. doi:10.1001/ab.21554

Watson, D. (2000). *Mood and temperament.* New York: Guildford Press.

Watson, D. (2005). Rethinking the mood and anxiety disorders: A quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*, *114*, 522-536. doi:10.1037/0021-843X.114.4.522

Watson, D. & Clark, L. A. (2011). Emotions and the emotional disorder: A quantitative hierarchical perspective. *International Journal of Clinical and Health Psychology*, *11*, 429-442. Retrieved from https://www.redalyc.org/pdf/337/33719289001.pdf.

Watson, D., Gamez, W., Simms, L. J. (2005). Basic dimensions of temperament and their relation to anxiety and depression: A symptom-based perspective. *Journal of Research in Personality*, *39*, 46–66. doi:10.1016/j.jrp.2004.09.006

Weathers, F. W. (2017). Redefining posttraumatic stress disorder for DSM-5. *Current Opinion in Psychology*, *14*, 122–126. doi:10.1016/j.copsyc.2017.01.002

Weathers, F.W., Blake, D.D., Schnurr, P.P., Kaloupek, D.G., Marx, B.P., & Keane, T.M. (2013b). *The Life Events Checklist for DSM-5 (LEC-5)*. Available from http://www.ptsd.va.gov

Weathers, F.W., Blake, D.D., Schnurr, P.P., Kaloupek, D.G., Marx, B.P., & Keane, T.M. (2013c). *The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5)*. Available from www.ptsd.va.gov

Weathers, F., Litz, B., Herman, D., Huska, J., & Keane, T. (October 1993). *The PTSD Checklist (PCL): Reliability, Validity, and Diagnostic Utility*. Paper presented at the Annual Convention of the International Society for Traumatic Stress Studies, San Antonio, TX.

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013a). *The PTSD Checklist for DSM-5 (PCL-5)*. Available from http://www.ptsd.va.gov

Weathers, F. W., Marx, B. P., Friedman, M. J., & Schnurr, P. P. (2014). Posttraumatic stress disorder in DSM-5: New criteria, new measures, and implications for assessment. *Psychological Injury and Law*, *7*(2), 93–107. doi:10.1007/s12207-014-9191-1

Wortmann, J. H., Jordan, A. H., Weathers, F. W., Resick, P. A., Dondanville, K. A., Hall-Clark, B., … Litz, B. T. (2016). Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. *Psychological Assessment*, *28*(11), 1392–1403. doi:10.1037/pas0000260

Yufik, T., & Simms, L. J. (2010). A meta-analytic investigation of the structure of posttraumatic stress disorder symptoms. *Journal of Abnormal Psychology*, *119*(4), 764–776. doi:10.1037/a0020981

Table 1

Item mappings for confirmatory factor analytic models of *DSM-IV* PTSD.

PCL Item	*DSM-IV*	Emotional Numbing	Dysphoria	Dysphoric Arousal
1: Memories	INT	INT	INT	INT
2: Dreams	INT	INT	INT	INT
3: Flashbacks	INT	INT	INT	INT
4: Cued distress	INT	INT	INT	INT
5: Cued physical reactions	INT	INT	INT	INT
6: Avoiding internal reminders	AV/N	AV	AV	AV
7: Avoiding external reminders	AV/N	AV	AV	AV
8: Amnesia	AV/N	N	D	N
9: Loss of interest	AV/N	N	D	N
10: Detachment or estrangement	AV/N	N	D	N
11: Numbing	AV/N	N	D	N
12: Hopelessness	AV/N	N	D	N
13: Sleep	AR	AR	D	DA
14: Irritability	AR	AR	D	DA
15: Concentration	AR	AR	D	DA
16: Hypervigilance	AR	AR	AR	AA
17: Startle	AR	AR	AR	AA

Note. *DSM-IV* = Diagnostic and Statistical Manual of Mental Disorders (4th ed.; APA, 1994); PCL = PTSD Checklist (Weathers, Herman, Huska, & Keane, 1993); INT = Intrusions; AV = Avoidance; N = Numbing; AR = Arousal; D = Dysphoria; DA = Dysphoric arousal; AA = Anxious arousal.

Table 2

Item mappings for confirmatory factor analytic models of *DSM-5* PTSD.

PTSD symptoms	*DSM-5*	Dysphoric Arousal	Anhedonia	Externalizing Behavior	Hybrid
1. Memories	REX	REX	REX	REX	REX
2. Dreams	REX	REX	REX	REX	REX
3. Flashbacks	REX	REX	REX	REX	REX
4. Cued distress	REX	REX	REX	REX	REX
5. Cued physical reactions	REX	REX	REX	REX	REX
6. Avoiding internal reminders	AV	AV	AV	AV	AV
7. Avoiding external reminders	AV	AV	AV	AV	AV
8. Amnesia	NACM	NACM	NACM	NACM	NA
9. Negative beliefs	NACM	NACM	NACM	NACM	NA
10. Blame	NACM	NACM	NACM	NACM	NA
11. Negative feelings	NACM	NACM	NACM	NACM	NA
12. Loss of interest	NACM	NACM	AN	NACM	AN
13. Detachment	NACM	NACM	AN	NACM	AN
14. Numbing	NACM	NACM	AN	NACM	AN
15. Irritability and aggression	AAR	DA	DA	EB	EB
16. Reckless behavior	AAR	DA	DA	EB	EB
17. Hypervigilance	AAR	AA	AA	AA	AA
18. Startle	AAR	AA	AA	AA	AA
19. Concentration	AAR	DA	DA	DA	DA
20. Sleep	AAR	DA	DA	DA	DA

Note. *DSM-5* = Diagnostic and Statistical Manual of Mental Disorders (5^{th} ed.; APA, 2013); REX = Reexperiencing; AV = Avoidance; NACM = Negative alterations in cognition and mood; AAR = Alterations in arousal and reactivity; Dysphoric arousal; AA = Anxious arousal; EB = Externalizing behavior.

Table 3

Descriptive statistics for scales used in analyses.

Variable	*N*	*M*	*SD*	Possible range	Observed range	α
PCL-5						
Total Scale	1114	15.83	17.26	0-80	0-80	.96
REX	1131	4.24	4.87	0-20	0-20	.91
AV	1138	2.26	2.52	0-8	0-8	.87
NACM	1129	5.19	6.47	0-28	0-28	.90
AAR	1133	4.12	5.35	0-24	0-24	.88
NA	1134	3.44	4.04	0-16	0-16	.83
AN	1134	1.75	2.94	0-12	0-12	.91
EB	1139	0.80	1.53	0-8	0-8	.71
AA	1135	1.62	2.25	0-8	0-8	.83
DA	1136	1.71	2.35	0-8	0-8	.86
PAI						
DEP	1038	20.55	12.52	0-72	0-65	.91
ANX	1040	26.53	14.44	0-72	0-72	.93
AGG	1051	13.88	9.09	0-54	0-49	.86
ANT	1036	16.38	9.59	0-72	0-56	.83
ALC	1062	6.13	5.81	0-36	0-35	.82
DRG	1067	5.31	5.26	0-36	0-32	.73

Note. PCL-5 = PTSD Checklist for *DSM-5*; REX = Reexperiencing; AV = Avoidance; NAMC = Negative alterations in cognitions and mood; AAR = Alterations in arousal and reactivity; NA = Negative affect; AN = Anhedonia; EB = Externalizing behavior; AA = Anxious arousal; DA = Dysphoric arousal; PAI = Personality Assessment Inventory; DEP = Depression; ANX = Anxiety; AGG = Aggression; ANT = Antisocial features; ALC = Alcohol problems; DRG = Drug problems.

Table 4

Item-level descriptive statistics for PCL-5 items.

Item	*N*	*M*	*SD*	Possible range	Observed range
1	1139	0.95	1.15	0-4	0-4
2	1139	0.58	1.03	0-4	0-4
3	1138	0.53	0.99	0-4	0-4
4	1135	1.32	1.29	0-4	0-4
5	1136	0.86	1.22	0-4	0-4
6	1138	1.24	1.35	0-4	0-4
7	1139	1.02	1.33	0-4	0-4
8	1139	0.74	1.10	0-4	0-4
9	1139	0.75	1.20	0-4	0-4
10	1137	0.90	1.32	0-4	0-4
11	1136	1.10	1.32	0-4	0-4
12	1139	0.51	1.03	0-4	0-4
13	1136	0.66	1.11	0-4	0-4
14	1137	0.58	1.06	0-4	0-4
15	1139	0.48	0.94	0-4	0-4
16	1139	0.32	0.79	0-4	0-4
17	1139	0.92	1.27	0-4	0-4
18	1135	0.70	1.16	0-4	0-4
19	1137	0.84	1.24	0-4	0-4
20	1138	0.87	1.28	0-4	0-4

Table 5

Fit statistics for models of PTSD symptom structure using the PCL-5.

Model	χ^2	*df*	*p*	RMSEA (90% CI)	*p*	SRMR	CFI	TLI	AIC	BIC
4-factor *DSM-5* CFA	1196.57	164	<.01	0.07 (0.07, 0.08)	.00	0.05	0.89	0.88	55,133.61	55,466.11
7-factor Hybrid CFA	524.14	149	<.01	0.05 (0.04, 0.05)	.87	0.03	0.96	0.95	53,933.74	54,341.81
4-factor DSM-5 ESEM	602.78	116	<.01	0.06 (0.05, 0.07)	.00	0.02	0.95	0.92	54,047.87	54,622.19
7-factor Hybrid ESEM	140.95	71	<.01	0.03 (0.02, 0.04)	1.00	0.01	0.99	0.98	53,396.41	54,197.44

Notes. *N* = 1139 for all models. PCL-5 = PTSD Checklist for *DSM-5*; *DSM-5* = *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition*; CFA = confirmatory factor analysis; ESEM = Exploratory structural equation model; *df* = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; CI = Confidence Interval; SRMR = Standardized Root Mean Square Residual; CFI = Bentler's Comparative Fit Index; TLI = Tucker Lewis Index; AIC = Akaike information criterion; BIC = Bayesian information criterion.

Table 6

Satorra-Bentler χ^2 difference test between 4- and 7- factor ESEM and CFA models.

$Model_1$	$Model_0$	C_1	C_0	Satorra-Bentler Scaled χ^2	*df*	*p*
7-factor ESEM	4-factor ESEM	1.52	1.59	439.00	45	<0.01
7-factor ESEM	7-factor CFA	1.52	1.73	360.33	78	<0.01
4-factor ESEM	4-factor CFA	1.59	1.79	520.23	48	<0.01
7-factor CFA	4-factor CFA	1.73	1.79	527.70	15	<0.01

Note. *N* = 1139. PCL-5 = PTSD Checklist for *DSM-5*; *DSM-5 = Diagnostic and Statistical Manual of Mental Disorders, 5th Edition*; $Model_1$ = freely-estimated model; $Model_0$ = constrained model; ESEM = Exploratory structural equation model; CFA = confirmatory factor analysis; C = scaling factor; *df* = degrees of freedom

Table 7
Standardized factor loadings for 4-factor PTSD models on the PCL-5.

	ESEM 4-factor				CFA 4-factor			
PCL-5 Item	REX	AV	NACM	AAR	REX	AV	NACM	AAR
1. Memories	**.87***	.02	.01	-.01	.86*			
2. Dreams	**.78***	-.05	-.08	.10	.79*			
3. Flashbacks	**.74***	-.03	-.14*	.22*†	.79*			
4. Cued distress	**.73***	.15†	.16*	-.06	.82*			
5. Physical reactions	**.67***	.13*†	.06	.14*†	.84*			
6. Internal avoidance	.46*†	**.18**	.33*†	.05		.89*		
7. External avoidance	.46*†	**.18**	.29*†	.09		.88*		
8. Amnesia	.07	.05	**.36***	.07			.46*	
9. Negative beliefs	-.13*	.14*†	**.84***	.13*†			.81*	
10. Blame	-.02	.35*†	**.86***	-.04			.71*	
11. Negative feelings	.12*	.28*†	**.71***	.07			.79*	
12. Loss of interest	.09	-.28*	**.48***	.24*†			.80*	
13. Detachment	.05	-.27*	**.66***	.16*†			.86*	
14. Numbing	.03	-.31*	**.65***	.16*†			.85*	
15. Angry reactions	.05	-.23*	.33*†	**.39***				.75*
16. Reckless behavior	.04	-.21*	.33*†	**.25***				.61*
17. Hypervigilance	-.01	.20*†	.05	**.73***				.69*
18. Startle	-.07*	.19†	-.26*	**1.20***				.78*
19. Concentration	.11	-.17*	.27*†	**.48***				.85*
20. Sleep	.18*	-.11*	.20*†	**.44***				.80*

Note. PTSD = Posttraumatic stress disorder; PCL-5 = PTSD Checklist for *DSM-5*; ESEM = Exploratory structural equation modeling; CFA = Confirmatory factor analysis; REX = Reexperiencing; AV = Avoidance; NAMC = Negative alterations in cognitions and mood; AAR = Alterations in arousal and reactivity.
*p < .05; †salient

Table 8

Standardized factor loadings for 7-factor PTSD models on the PCL-5.

	ESEM 7-factor							CFA 7-factor						
PCL-5 Item	REX	AV	NA	AN	EB	AA	DA	REX	AV	NA	AN	EB	AA	DA
1	**.91***	-.06	.06	-.05	.01	-.11*	.08	.86*						
2	**.86***	.00	-.02	-.12*	.28*†	-.07*	.06	.79*						
3	**.82***	-.07	-.05	.05	.09	.14*	-.07	.79*						
4	**.69***	.18*	.01	.09	-.34*	-.05	.04	.82*						
5	**.68***	.10	.04	.09	-.13*	.15*	-.11	.84*						
6	.08	**.74***	.06	-.04	-.02	.00	.20*		.89*					
7	.10	**.76***	.04	-.02	.07	.07	-.04		.88*					
8	-.07	.27*	**.18***	.05	.08	.00	.10			.46*				
9	-.07	.02	**.68***	.20*	.03	.07	.02			.82*				
10	-.04	-.02	**1.02***	-.11*	.03	-.02	-.06			.82*				
11	.17*	-.04	**.73***	.02	-.10*	.04	.03			.89*				
12	.16*	.01	.08	**.52***	.15*†	.08	.07				.83*			
13	.08*	.06	.17*	**.56***	.07	-.01	.17*				.90*			
14	.05	.05	.15*	**.55***	.12*	-.04	.21*				.89*			
15	.07	.09	-.01	.44*†	**.14**	.22*	.08					.84*		
16	.07	.08	.08	.32*†	**.28***	.11	.02					.67*		
17	.00	.14*	.14*	-.04	.02	**.61***	.00						.79*	
18	.06	-.04	.00	-.11*	.03	**.88***	.14*						.90*	
19	.00	.04	.00	.14	-.03	.12	**.69***							.89*
20	.07	.01	.04	-.02	-.01	.07	**.74***							.84*

Note. PTSD = Posttraumatic stress disorder; PCL-5 = PTSD Checklist for *DSM-5*; ESEM = Exploratory structural equation modeling; CFA = Confirmatory factor analysis; REX = Reexperiencing; AV = Avoidance; NA = Negative affect; AN = Anhedonia; EB = Externalizing behavior; AA = Anxious arousal; DA = Dysphoric arousal. Anchor items for each factor appear in bold.
*$p < .05$; †salient

Table 9

Standardized factor intercorrelations of 7-factor PTSD models on the PCL-5.

	ESEM 7-factor						CFA 7-factor					
	AV	NA	AN	EB	AA	DA	AV	NA	AN	EB	AA	DA
REX	.74*	.76*	.45*	.16*	.64*	.67*	.85*	.81*	.69*	.67*	.73*	.73*
AV		.71*	.43*	-.02	.54*	.54*		.80*	.66*	.64*	.68*	.68*
NA			.57*	.16*	.57*	.65*			.80*	.72*	.69*	.74*
AN				.26*	.47*	.64*				.91*	.66*	.86*
EB					.19*	.27*					.74*	.82*
AA						.66*						.80*
	M = .48						*M* = .75					

Note. PCL-5 = PTSD Checklist for *DSM-5*; ESEM = Exploratory structural equation modeling; CFA = Confirmatory factor analysis; REX = Reexperiencing; AV = Avoidance; NAMC = Negative alterations in cognitions and mood; AAR = Alterations in arousal and reactivity; NA = Negative affect; AN = Anhedonia; EB = Externalizing behavior; AA = Anxious arousal; DA = Dysphoric arousal

*p < .05.

Table 10

Standardized within- and between-cluster correlations of 2 PCL-5 and PAI factors.

	1. PAI-Depression	2. PAI-Anxiety	3. PAI-Externalizing Behavior	Within-cluster Wald W (Cohen's q)		
				$W\ (q)_{2\text{-}1}$	$W\ (q)_{3\text{-}1}$	$W\ (q)_{3\text{-}2}$
CFA						
PCL-Negative Affect	.52	.50	.20	0.00 (.03)	6.91* (.38)	5.97* (.35)
PCL-Anhedonia	.63	.49	.27	12.43* (.21)	4.15 (.46)	0.26 (.25)
Between-cluster Wald W (Cohen's q)	48.45* (.16)	23.18* (.01)	24.97* (.08)			
ESEM						
PCL-Negative Affect	.49	.46	.18	0.06 (.04)	6.99* (.35)	5.40* (.32)
PCL-Anhedonia	.60	.45	.21	12.40* (.21)	7.70* (.48)	1.39 (.27)
Between-cluster Wald W (Cohen's q)	6.92* (.15)	0.11 (.02)	0.55 (.03)			

Note. PCL-5 = Posttraumatic Stress Disorder Checklist for *DSM-5*; PAI = Personality Assessment Inventory; Parentheses indicate pairwise Cohen's q effect size magnitude.

*Statistically significant Wald W at $p < .05$ after applying Benjamini and Hochberg's (1995) procedure to control for the false discovery rate.

Table 11

Standardized within- and between-cluster correlations of 3 PCL-5 and PAI factors.

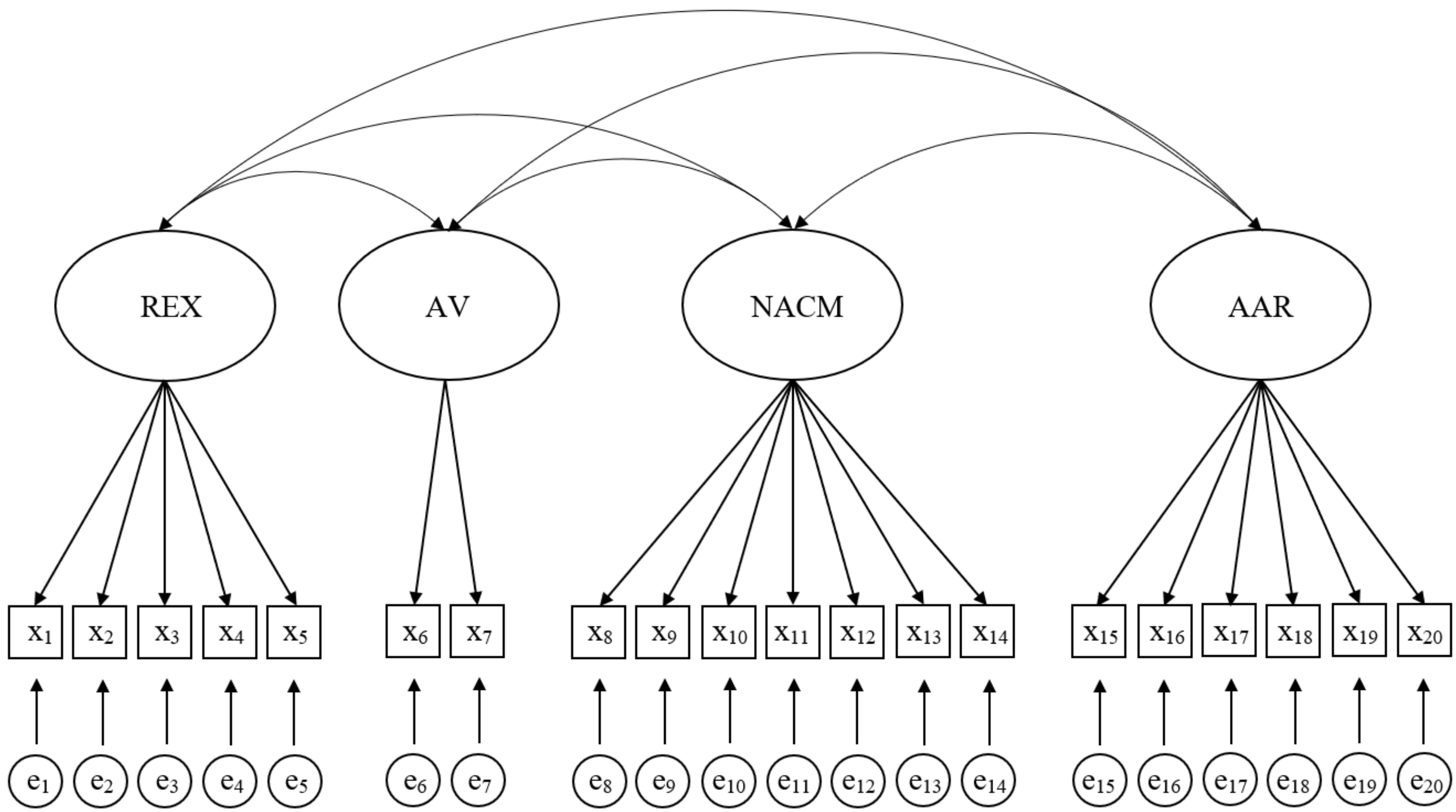
	1. PAI-Depression	2. PAI-Anxiety	3. PAI-Externalizing Behavior	Within-cluster Wald W (Cohen's q)		
				W $(q)_{2\text{-}1}$	W $(q)_{3\text{-}1}$	W $(q)_{3\text{-}2}$
CFA						
A. PCL-Externalizing Behavior	.56	.48	.42	3.89 (.11)	3.02 (.19)	4.99* (.08)
B. PCL-Anxious Arousal	.45	.54	.19	13.46* (.12)	3.29 (.29)	9.02* (.41)
C. PCL-Dysphoric Arousal	.62	.57	.27	0.35 (.08)	4.41 (.45)	2.81 (.37)
Between-cluster Wald W (Cohen's $q)_{b\text{-}a}$	0.40 (.15)	25.25* (.09)	18.49* (.25)			
Between-cluster Wald W (Cohen's $q)_{c\text{-}a}$	40.73* (.09)	48.68* (.13)	0.373 (.17)			
Between-cluster Wald W (Cohen's $q)_{c\text{-}b}$	35.51* (.23)	6.60* (.04)	12.02* (.02)			
ESEM						
A. PCL-Externalizing Behavior	.40	.24	.51	9.87* (.18)	20.84* (.14)	25.43* (.32)
B. PCL-Anxious Arousal	.41	.52	.16	14.95* (.14)	3.37 (.28)	9.84* (.41)
C. PCL-Dysphoric Arousal	.61	.54	.27	0.95 (.11)	4.01 (.44)	1.54 (.33)
Between-cluster Wald W (Cohen's $q)_{b\text{-}a}$	0.026 (.01)	9.96* (.33)	29.05* (.40)			
Between-cluster Wald W (Cohen's $q)_{c\text{-}a}$	11.01* (.29)	13.73* (.37)	17.06* (.28)			
Between-cluster Wald W (Cohen's $q)_{c\text{-}b}$	21.76* (.28)	0.269 (.04)	5.89* (.11)			

Note. PCL-5 = Posttraumatic Stress Disorder Checklist for *DSM-5*; PAI = Personality Assessment Inventory; Parentheses indicate pairwise Cohen's q effect size magnitude.

*Statistically significant Wald W at $p < .05$ after applying Benjamini and Hochberg's (1995) procedure to control for the false discovery rate.

Figure 1

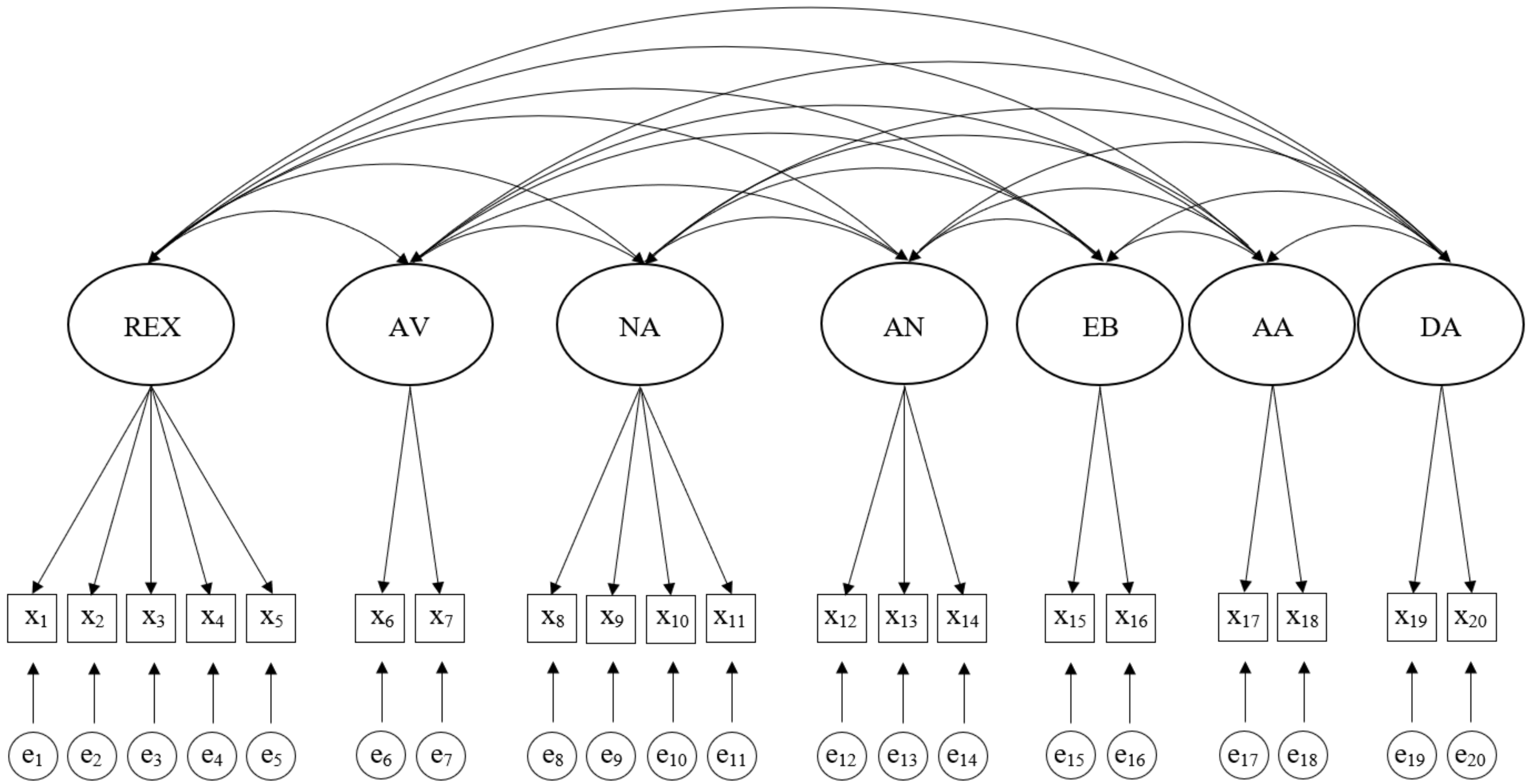
Model structure of *DSM-5* PTSD in CFA.



Note. X_1-X_{20} are PTSD symptoms 1-20; e_1-e_{20} are observed indicators PCL-5 items 1-20; REX Reexperiencing; AV = Avoidance; NACM = Negative alterations in cognitions and mood; AAR = Alterations in arousal and reactivity.

Figure 2

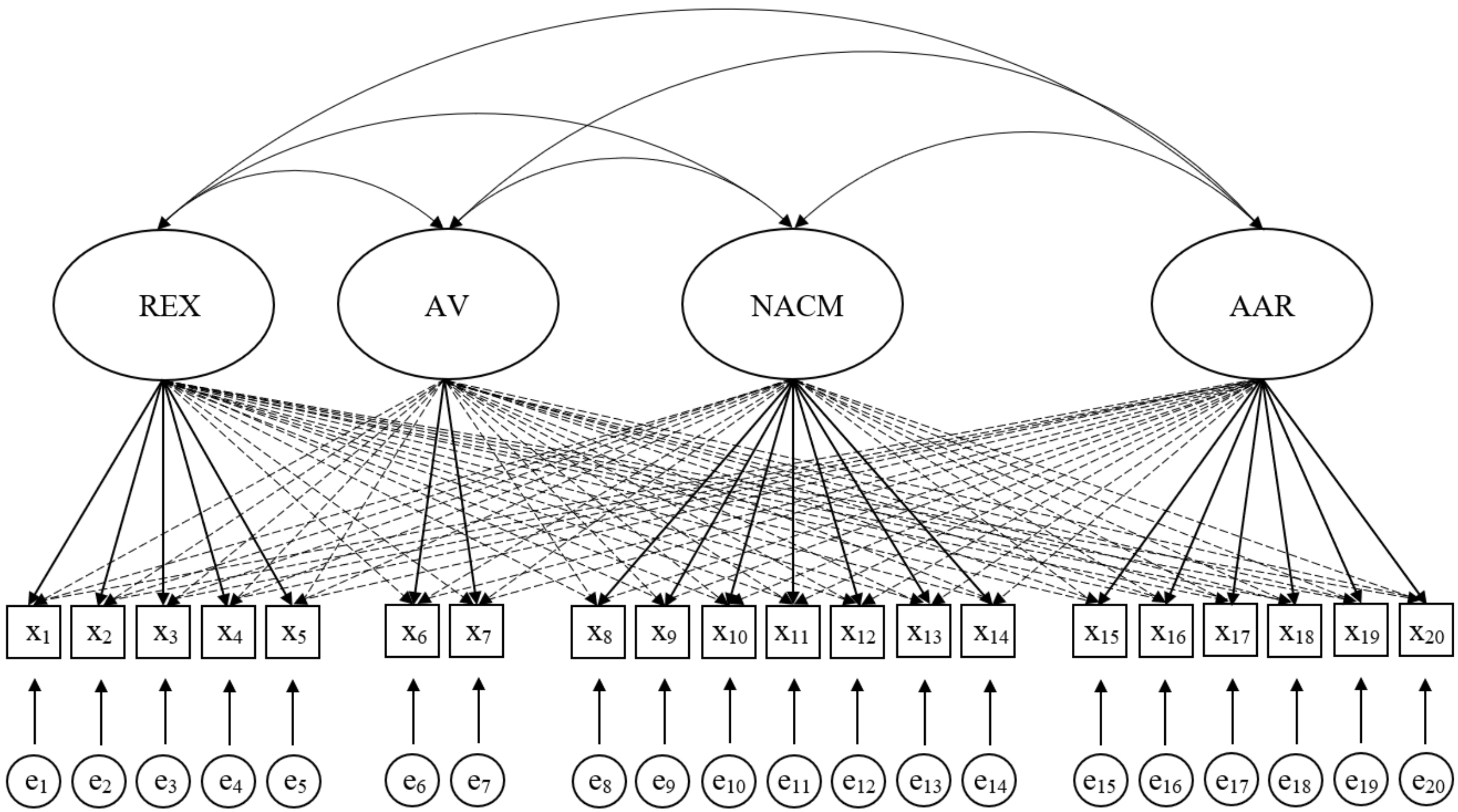
Model structure of the 7-factor Hybrid model of PTSD in CFA.



Note. X_1-X_{20} are PTSD symptoms 1-20; e_1-e_{20} are observed indicators PCL-5 items 1-20; REX Reexperiencing; AV = Avoidance; NA = Negative affect; AN = Anhedonia; EB = Externalizing behavior; AA = Anxious Arousal; DA = Dysphoric Arousal

Figure 3

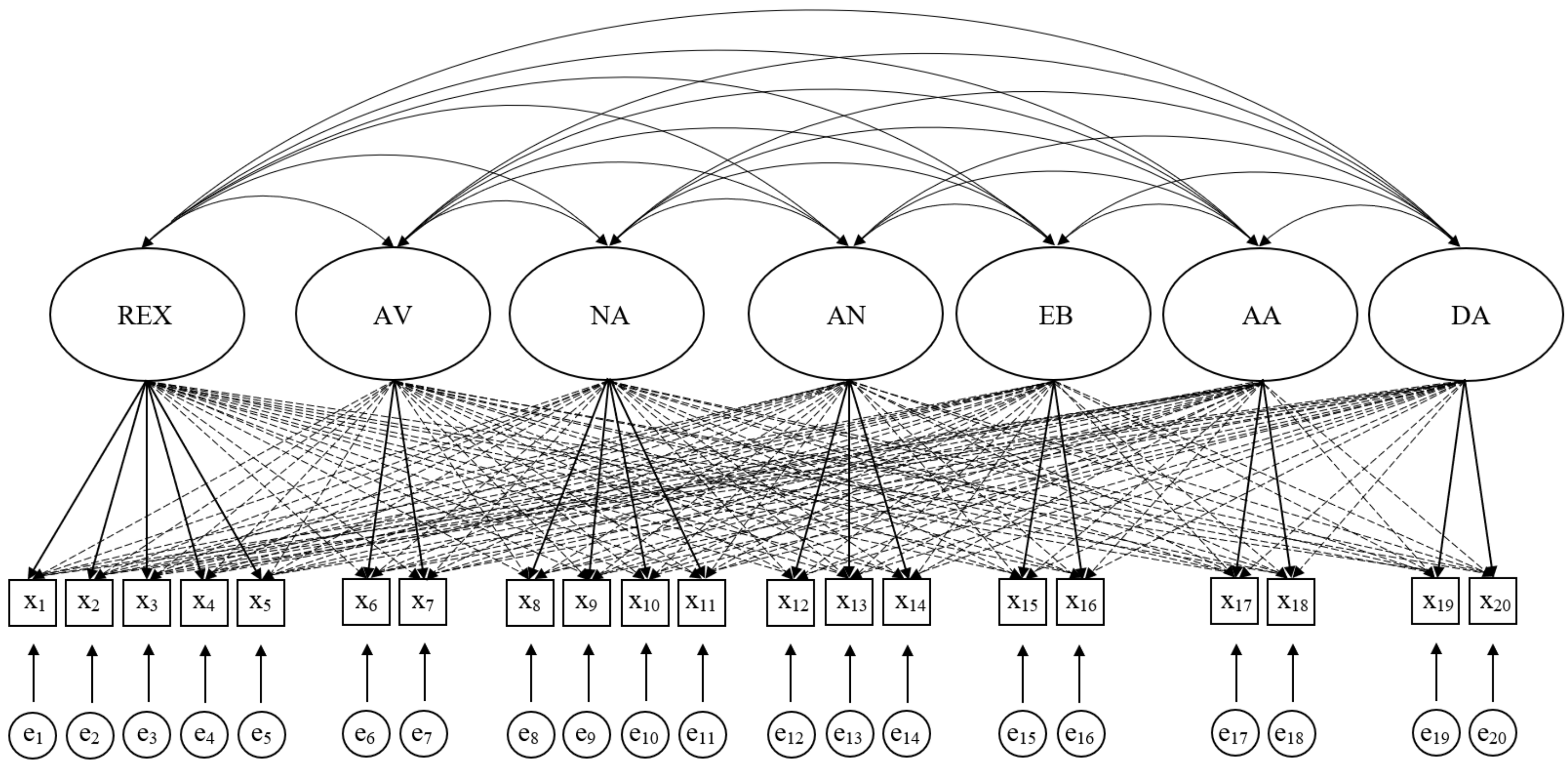
Model structure of *DSM-5* PTSD in ESEM.



Note. X_1-X_{20} are PTSD symptoms 1-20; e_1-e_{20} are observed indicators PCL-5 items 1-20; REX Reexperiencing; AV = Avoidance; NACM = Negative alterations in cognitions and mood; AAR = Alterations in arousal and reactivity; Solid lines represent primary factor loadings; Dotted lines represent secondary factor loadings (cross-loadings)

Figure 4

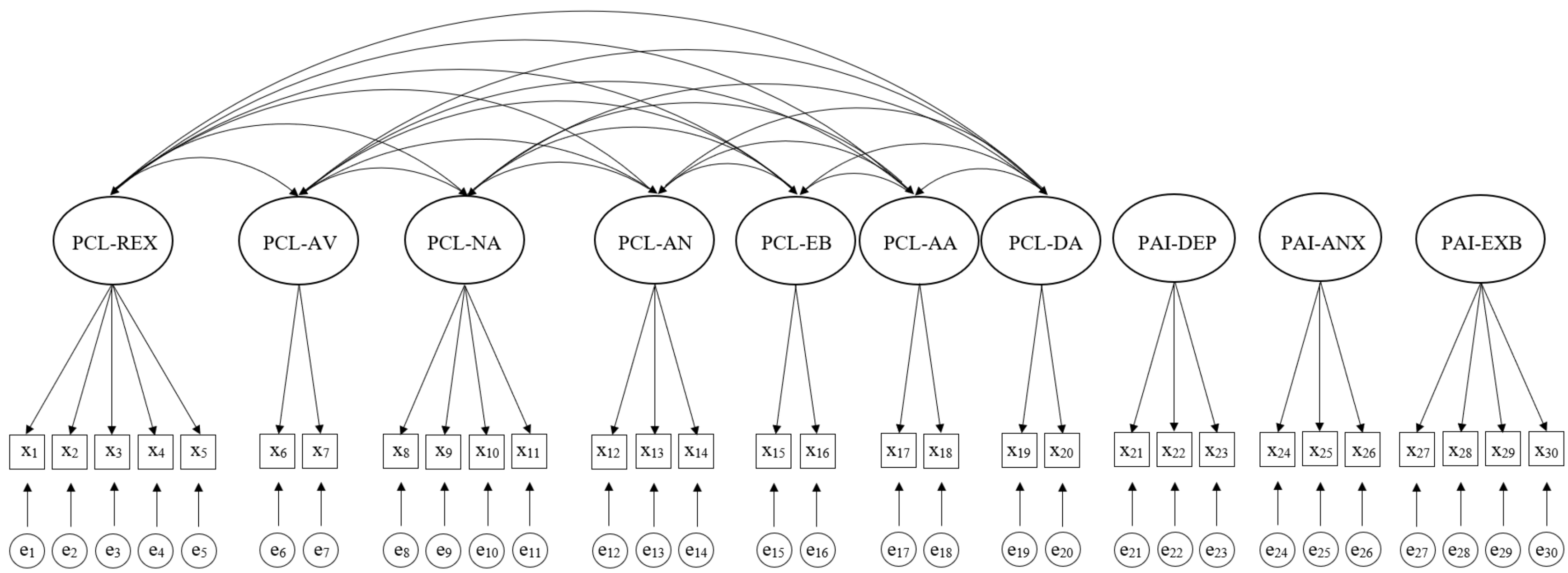
Model structure of the 7-factor Hybrid model of PTSD in ESEM.



Note. X_1-X_{20} are PTSD symptoms 1-20; e_1-e_{20} are observed indicators PCL-5 items 1-20; REX Reexperiencing; AV = Avoidance; NA = Negative affect; AN = Anhedonia; EB = Externalizing behavior; AA = Anxious Arousal; DA = Dysphoric Arousal; Solid lines represent primary factor loadings; Dotted lines represent secondary factor loadings (cross-loadings).

Figure 5

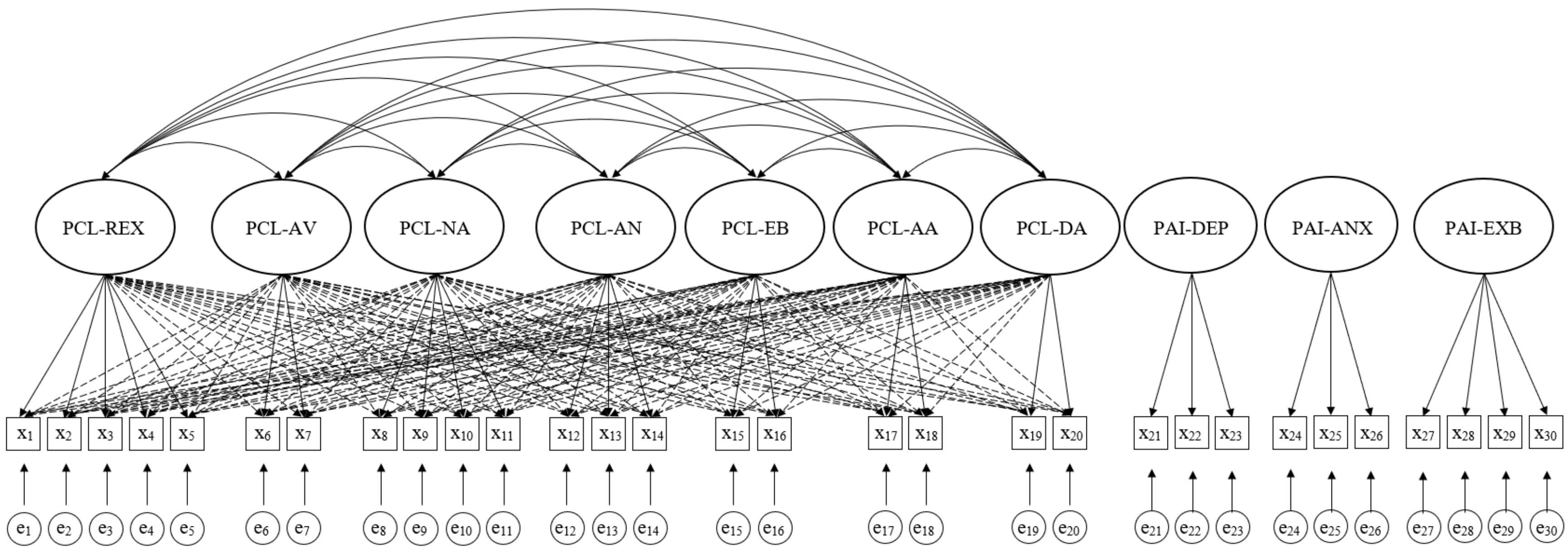
Model structure of 10-factor CFA including PCL-5 and PAI latent variables.



Note. X_1-X_{20} are PTSD symptoms 1-20; X_{21}-X_{30} are symptoms of depression, anxiety, and externalizing behavior; e_1-e_{20} are PCL-5 items 1-20; e_{21} = PAI Depression Cognitive subscale (DEP-C); e_{22} = PAI Depression Affective subscale (DEP-A); e_{23} = PAI Depression Physiological subscale (DEP-P); e_{24} = PAI Anxiety Cognitive subscale (ANX-C); e_{25} = PAI Anxiety Affective subscale (ANX-A); e_{26} = PAI Anxiety Physiological subscale (ANX-P); e_{27} = PAI Antisocial Features scale (ANT); e_{28} = PAI Aggression scale (AGG); e_{29} = PAI Anxiety Alcohol Problems scale (ALC); e_{30} = PAI Drug Problems scale (DRG); PCL-REX Reexperiencing; PCL-AV = Avoidance; PCL-NACM = Negative alterations in cognitions and mood; PCL-AAR = Alterations in arousal and reactivity; PAI-DEP = Depression; PAI-ANX = Anxiety; PAI-EXB = Externalizing Behavior.

Figure 6

Model structure of 10-factor ESEM including PCL-5 and PAI latent variables.



Note. X_1-X_{20} are PTSD symptoms 1-20; X_{21}-X_{30} are symptoms of depression, anxiety, and externalizing behavior; e_1-e_{20} are PCL-5 items 1-20; e_{21} = PAI Depression Cognitive subscale (DEP-C); e_{22} = PAI Depression Affective subscale (DEP-A); e_{23} = PAI Depression Physiological subscale (DEP-P); e_{24} = PAI Anxiety Cognitive subscale (ANX-C); e_{25} = PAI Anxiety Affective subscale (ANX-A); e_{26} = PAI Anxiety Physiological subscale (ANX-P); e_{27} = PAI Antisocial Features scale (ANT); e_{28} = PAI Aggression scale (AGG); e_{29} = PAI Anxiety Alcohol Problems scale (ALC); e_{30} = PAI Drug Problems scale (DRG); PCL- REX Reexperiencing; PCL-AV = Avoidance; PCL-NA = Negative affect; PCL-AN = Anhedonia; PCL-EB = Externalizing behavior; PCL-AA = Anxious Arousal; PCL-DA = Dysphoric Arousal; PAI-DEP = Depression; PAI-ANX = Anxiety; PAI-EXB = Externalizing Behavior; Solid lines represent primary factor loadings; Dotted lines represent secondary factor loadings (cross-loadings).

Appendix A

```
TITLE: 7-factor Hybrid ESEM on PCL-5;

DATA: FILE IS dataset.dat;

VARIABLE:
NAMES ARE ID PCL1 PCL2 PCL3 PCL4 PCL5 PCL6 PCL7 PCL8 PCL9 PCL10
PCL11 PCL12 PCL13 PCL14 PCL15 PCL16 PCL17 PCL18 PCL19 PCL20;
USEVARIABLES ARE PCL1 PCL2 PCL3 PCL4 PCL5 PCL6 PCL7 PCL8 PCL9 PCL10
PCL11 PCL12 PCL13 PCL14 PCL15 PCL16 PCL17 PCL18 PCL19 PCL20;
IDVARIABLE=ID;
MISSING ARE ALL (-999);

ANALYSIS:
ESTIMATOR IS MLR;
ROTATION IS TARGET;

MODEL:
intrus by PCL1 PCL2 PCL3 PCL4 PCL5
PCL6~0 PCL7~0 PCL8~0 PCL9~0 PCL10~0 PCL11~0 PCL12~0 PCL13~0
PCL14~0 PCL15~0 PCL16~0 PCL17~0 PCL18~0 PCL19~0 PCL20~0 (*1);

avoid by PCL6 PCL7
PCL1~0 PCL2~0 PCL3~0 PCL4~0 PCL5~0 PCL8~0 PCL9~0 PCL10~0 PCL11~0 PCL12~0
PCL13~0 PCL14~0 PCL15~0 PCL16~0 PCL17~0 PCL18~0 PCL19~0 PCL20~0 (*1);

naegaf by PCL8 PCL9 PCL10 PCL11
PCL1~0 PCL2~0 PCL3~0 PCL4~0 PCL5~0 PCL6~0 PCL7~0 PCL12~0 PCL13~0 PCL14~0
PCL15~0 PCL16~0 PCL17~0 PCL18~0 PCL19~0 PCL20~0 (*1);

anhed by PCL12 PCL13 PCL14
PCL1~0 PCL2~0 PCL3~0 PCL4~0 PCL5~0 PCL6~0 PCL7~0 PCL8~0 PCL9~0 PCL10~0
PCL11~0 PCL15~0 PCL16~0 PCL17~0 PCL18~0 PCL19~0 PCL20~0 (*1);

extbeh by PCL15 PCL16
PCL1~0 PCL2~0 PCL3~0 PCL4~0 PCL5~0 PCL6~0 PCL7~0 PCL8~0 PCL9~0 PCL10~0
PCL11~0 PCL12~0 PCL13~0 PCL14~0 PCL17~0 PCL18~0 PCL19~0 PCL20~0 (*1);

anxar by PCL17 PCL18
PCL1~0 PCL2~0 PCL3~0 PCL4~0 PCL5~0 PCL6~0 PCL7~0 PCL8~0 PCL9~0 PCL10~0
PCL11~0 PCL12~0 PCL13~0 PCL14~0 PCL15~0 PCL16~0 PCL19~0 PCL20~0 (*1);

dysar by PCL19 PCL20
PCL1~0 PCL2~0 PCL3~0 PCL4~0 PCL5~0 PCL6~0 PCL7~0 PCL8~0 PCL9~0 PCL10~0
PCL11~0 PCL12~0 PCL13~0 PCL14~0 PCL15~0 PCL16~0 PCL17~0 PCL18~0 (*1);

OUTPUT: SAMPSTAT STANDARDIZED MODINDICES tech1 tech4;
```