Posttraumatic stress disorder: An empirical evaluation of core assumptions

Gerald M. Rosen a,*, Scott O. Lilienfeld b

a University of Washington, Seattle, Washington, United States
b Emory University, Atlanta, Georgia, United States

Received 2 April 2007; received in revised form 15 December 2007; accepted 17 December 2007

Abstract

The diagnosis of posttraumatic stress disorder (PTSD) rests on several core assumptions, particularly the premise that a distinct class of traumatic events is linked to a distinct clinical syndrome. This core assumption of specific etiology ostensibly distinguishes the PTSD diagnosis from virtually all other psychiatric disorders. Additional attempts to distinguish PTSD from extant conditions have included searches for distinctive markers (e.g., biological and laboratory findings) and hypothesized underlying mechanisms (e.g., fragmentation of traumatic memory). We review the literature on PTSD’s core assumptions and various attempts to validate the construct within a nomological network of distinctive correlates. We find that virtually all core assumptions and hypothesized mechanisms lack compelling or consistent empirical support. We consider the implications of these findings for conceptualizing PTSD in the forthcoming edition of the American Psychiatric Association’s diagnostic manual.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Posttraumatic stress disorder (PTSD); Validity; Construct validity; Discriminant validity; Traumatic events; Criterion A; Symptom criteria

Contents

1. Assumption of a specific etiology
   1.1. Criterion A as a necessary condition
   1.2. The adversity–stress model and assumption of greatest influence
   1.3. Changing definitions, A2, and criterion creep

2. Is PTSD a distinct clinical syndrome?
   2.1. Variability of symptom patterns
   2.2. Comorbidity
   2.3. Emotional numbing
   2.4. The matter of bereavement

* Corresponding author. 117 East Louisa St., PMB-229, Seattle, WA. 98102, United States.
E-mail address: grosen@u.washington.edu (G.M. Rosen).

0272-7358/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
The construct validity of posttraumatic stress disorder (PTSD) has been a source of scientific controversy and debate ever since the diagnosis was introduced (DSM-III; American Psychiatric Association [APA], 1980). As observed by Spitzer, First, and Wakefield (2007): “Since its introduction into DSM-III in 1980, no other DSM diagnosis, with the exception of Dissociative Identity Disorder (a related disorder), has generated so much controversy in the field as to the boundaries of the disorder, diagnostic criteria, central assumptions, clinical utility, and prevalence in various populations (p. 233).” It is in this context that we take stock of the construct validity of the diagnosis of PTSD by critically examining its core assumptions and surrounding nomological network of falsifiable predictions (Cronbach & Meehl, 1955; Garber & Strassberg, 1991).

Before proceeding, we want to state clearly that we are not arguing that the intense psychological distress experienced by PTSD individuals is imaginary, or that the PTSD diagnosis is predictively useless. We fully acknowledge the substantial heuristic value of the PTSD construct in generating research on reactions to stress. Further, we acknowledge that studies comparing PTSD with non-PTSD individuals yield at least some clinically and statistically significant differences on important outcome measures, such as psychological dysfunction and physical disability (e.g. Norman, Stein & Davidson, 2007; Schnurr & Green, 2004; Zatzick et al., 1997). Given the serious and often disabling anxiety symptoms experienced by individuals who carry the PTSD diagnosis, such research findings are to be expected.

At the same time, it is appropriate — even necessary — to ask the critical question of whether PTSD, as a hypothetical construct (MacCorquodale & Meehl, 1948; Morey, 1991), is the best means of “carving nature at its joints” (Gangestad & Snyder, 1985) within the domain of anxiety disorders. That is, even if the PTSD diagnosis captures a clinically important grouping of individuals with severe psychological distress, it may not be the optimal category for conceptualizing or classifying such distress. It may also not possess substantial incremental validity (Sechrest, 1963) for predicting clinically important external validating criteria, above and beyond extant and better validated diagnoses (e.g., specific phobia, generalized anxiety disorder, depression).

The nomological network surrounding PTSD incorporates predictions regarding a variety of external validating criteria, including (a) the specificity of precipitating events, (b) relations between precipitating events and clinical symptoms, (c) discriminant validity (absence of excessive comorbidity) from other conditions, (d) psychophysiological reactivity, (e) neuroendocrine and brain imaging findings, and (f) distinctive features of traumatic memory. These nodes in the PTSD nomological network have assumed particular importance in the traumatology literature given that each could advance the claim that PTSD is distinctive from extant conditions, especially other anxiety disorders. In this paper, we (a) present the logical justification for each validating criterion’s inclusion in the PTSD nomological network, and (b) consider the extent to which each of these six linkages within the network have been either supported or called into question. By providing a comprehensive review of research bearing on each of these linkages we will evaluate systematically the construct validity of the PTSD diagnosis. We also will consider how a recent and influential attempt to evaluate the construct validity of PTSD, issued by the Institute of Medicine (IOM, 2006), failed to provide a balanced and scientifically informed presentation of the literature. Finally, we will consider proposals for the future of PTSD in the DSM-V, and offer recommendations and cautionary guidelines to address current controversies.
1. Assumption of a specific etiology

Since its inception, PTSD has rested on a core assumption: namely, that a distinct class of events (Criterion A: the "stressor criterion") is causally linked to a distinct set of reactions (Criteria B through D: the "symptom criteria"). The stressor criterion was intended to serve a "gatekeeper" function (Davidson & Foa, 1991), such that an individual could not receive a PTSD diagnosis without the occurrence of a traumatic event. In this respect, PTSD differs from virtually all other diagnoses in the DSM (e.g., schizophrenia, major depression, panic disorder) in that it is not agnostic with respect to etiology (Spitzer, 2001; Wilson, 1993). As a consequence, the PTSD diagnosis assumes a burden of proof not shared by other diagnoses: namely, a clear link between a precipitating stressor and resulting signs and symptoms.

Although Criterion A events are viewed by DSM-IV as necessary for the development of PTSD, they are not necessarily sufficient. This is because the most severe of stressors do not always produce the disorder. Andreasen (1980) provided an analogy by comparing the role of a stressor in PTSD with that of a force sufficient to break a leg. Andreasen observed that it can be normal for a leg to break when enough force is applied, although individual legs vary in the force required, time to heal, and degree of residual pathology. Andreasen’s analysis calls to mind Meehl’s (1977) classic discussion of specific etiology. Meehl observed that the strongest case for specific etiology, one in which there is no dispute over the use of that expression, is when the presence of agent \( x \) leads to a disease in all instances, and the disease cannot arise in the absence of \( x \). Meehl’s analysis allowed for weaker forms of causal specificity, such as when \( x \) exerts the greatest influence on the probability of disease but other factors play a role. Meehl called this latter situation the “uniformly most potent” kind of specific etiology.

Although the definition of Criterion A has changed in subsequent editions of the DSM (DSM-III-R: APA, 1987; DSM-IV: APA, 1994), an issue addressed later in this paper, the assumption has remained that traumatic events provide the uniformly most potent etiologic agent to account for PTSD. The adequacy of this assumption can be assessed by considering research on whether Criterion A events (a) are necessary for the development of PTSD’s defining symptom criteria, and (b) contribute the greatest variance to psychiatric morbidity.

1.1. Criterion A as a necessary condition

Not long after the creation of PTSD, Breslau and Davis (1987) challenged the validity of explicitly linking “a distinct symptomatic configuration with a distinct class of stressors” (p. 255). They noted that the specific stress-symptom connection hypothesized for PTSD, and purportedly observed by clinicians, had received little support from epidemiologic studies.

Numerous publications have supported Breslau and Davis’s challenge by demonstrating an association between non-Criterion A events and the full range of PTSD reactions. Research and clinical reports have documented that some individuals meet PTSD symptom criteria following such non-Criterion A events as marital disruption, affairs, and divorce (Burstein, 1995; Dattilio, 2004; Dreman, 1991; Helzer, Robins & McEvoy, 1987); collapse of adoption arrangements (Burstein, 1995); employment related stressors and money problems (Ravin & Boal, 1989; Scott & Stradling, 1994; Solomon & Canino, 1990); bereavement (Zisook, Chentsova-Dutton & Shuchter, 1998); childbirth (Ayers & Pickering, 2001; Czarnocka & Slade, 2000; Olde, van der Hart, Kleber & van Son, 2006); loss of cattle to foot and mouth disease (Olff, Koeter, Van Haaften, Kersten & Gersons, 2005); frightening Halloween television programs (Simons & Silveira, 1994); and breaking up with a best friend (Solomon & Canino, 1990). These findings have emerged across both various editions of the DSM (APA, 1980, 1987, 1994) and changing definitions of Criterion A.

Additional studies question the assumption that a subset of stressors is uniquely associated with PTSD. Gold, Marx, Soler-Baillo, and Sloan (2005) and Mol et al. (2005) conducted surveys among non-clinical populations and found that subjects who had not experienced a Criterion A event, reported reactions on a PTSD symptom checklist with reference to other life stressors (e.g., parental divorce, relationship problems, arrest) that were as high or higher as reactions reported by subjects who experienced trauma. It should be noted, however, that these studies did not establish a diagnosis of PTSD independent of symptom checklists. Although straightforward symptom checklists can perform well, in the sense of being validated by structured clinical interview (Brewin, 2005), these measures may confuse normal reactions with symptoms of disorder (Lees-Haley, Price, Williams & Betz, 2001). Therefore, the findings of Gold et al. and Mol et al. could reflect the endorsement of checklist items without confirming psychopathology. Nevertheless, the studies join with others to call into question the uniqueness of Criterion A events.
Bodkin, Pope, Detke, and Hudson (2007) administered the PTSD module of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID; First, Spitzer, Gibbon & Williams, 1996) to 103 sequential patients presenting for pharmacologic treatment of depression and anxiety, but not for PTSD per se. Patients who denied any history of trauma were asked whether there was a period of time when they were “plagued by distressing thoughts, worries, or fears.” All patients who did not report a Criterion A event were able to recall a distressing time, and they were instructed to refer to that period when responding to assessment questions from the SCID. Two blinded raters subsequently classified reported trauma histories or troubling thoughts as either traumatic or non-traumatic following the DSM-IV definition for Criterion A. Thirteen subjects had experiences that were judged equivocal (i.e., judges did not agree), leaving 54 clear-cut trauma and 36 clear-cut non-trauma subjects. Despite differences in the occurrence of a Criterion A event, 78% of patients in both groups met PTSD symptom criteria. Although the retrospective nature of this design limits conclusions, Bodkin et al. correctly observed that the occurrence of PTSD symptoms independent of trauma history calls into question the critical hypothesis of specific etiology upon which the diagnosis hinges. The import of Bodkin et al.’s findings is strengthened by a recent study on individuals with social anxiety disorder (Erwin, Heimberg, Marx & Franklin, 2006). In this research, participants with social anxiety disorder endorsed symptoms and fulfilled the clinical syndrome of PTSD with reference to uncomfortable social experiences that did not satisfy Criterion A. Taken together, the studies by Bodkin et al. and Erwin et al. raise significant questions. They indicate that the full clinical syndrome of PTSD can arise frequently among psychiatrically distressed subjects without any occurrence of a Criterion A event.

1.2. The adversity–stress model and assumption of greatest influence

Underlying the DSM’s concept of a distinct class of traumatic stressors is the concept of a dose–response relationship between stressor and symptoms. Because Criterion A events are by definition of high magnitude it was assumed they would be the greatest contributor to outcome (Andreasen, 1980). This assumption was basic to what has come to be called the adversity–stress model (Bowman & Yehuda, 2004).

The presence of a dose–response relationship, or “biological gradient,” is among Austin Bradford Hill’s (1965) famous criteria for inferring causality in epidemiology. As Hill noted, the absence of a dose–response relationship does not exclude a causal association, because the association may be of a more complex (e.g., curvilinear) form. Nevertheless, the existence of a dose–response relationship strengthens the argument for a causal association between stressors and criteria. Moreover, the absence of such a relation requires one to “envisage some much more complex relationship to satisfy the cause and effect hypothesis” (Hill, 1965, p. 296).

There now exists a substantial and complex literature on dose–response relationships, the adversity–stress model, and risk factors for PTSD. Several reviews (Bowman, 1997, 1999; Bowman & Yehuda, 2004; Breslau, 2002; Davidson & Fairbank, 1993; Gibbs, 1989; Lundin, 1995; McFarlane & de Girolamo, 1996; Ozer & Weiss, 2004), an edited text (Yehuda, 1999), structural equation modeling studies (King, King, Foy & Gudanowski, 1996; King, King, Foy, Keane & Fairbank, 1999; King, King, Fairbank, Keane & Adams, 1998a), and two meta-analyses (Brewin, Andrews & Valentine, 2000; Ozer, Best, Lipsey & Weiss, 2003) provide an overview of the issues. These sources, and research upon which they rely, find that (a) most individuals do not develop PTSD after Criterion A events, (b) a simple dose–response relationship is often not supported, and (c) factors extraneous to the event contribute more variance to clinical outcome than the event itself.

Epidemiologic studies (Breslau, Davis, Andreski & Peterson, 1991; Breslau et al., 1998; Davidson, Hughes, Blazer & George, 1991; Helzer et al., 1987; Kulka et al., 1990; Norris, 1992; Stein, Walker, Hazen & Forde, 1997b) demonstrate that the majority of adults encounter one or more Criterion A events during their lifetimes. Yet only a minority of individuals typically react to Criterion A events with symptoms of sufficient breadth, severity, and duration to meet the symptom criteria for PTSD. For example, in the field trial for DSM-IV, 93% of a community sample reported experiencing a Criterion A event, whereas only 10.3% met criteria for lifetime PTSD (Kilpatrick et al., 1998).

On occasion, higher rates of PTSD arise from the specific assessment methods used by researchers. The National Vietnam Veterans Readjustment Study (NVVRS; Kulka et al., 1990) caused considerable debate after finding that 30.9% of those serving in Vietnam developed PTSD, when only about 15% of Vietnam veterans had served in combat units (Shephard, 2001).Recently, Dohrenwend et al. (2006) reanalyzed data from the NVVRS using refined measures and taking into account ratings of impairment. This effort resulted in a downward adjustment of fully 40%, such that only 18.7% of veterans developed war-related PTSD during their lifetimes. In other instances, high rates of PTSD are
associated with the use of selected samples. In an unpublished study cited by Resnick, Kilpatrick, and Lipovsky (1991), 78% of victims referred by a rape crisis center reported reactions that met PTSD criteria. Phillips, Rosen, Zoellner, and Feeny (2006) reported that 82% of 17 Malaysian women in a shelter for battered women met PTSD criteria on the PSS-SR (Foa, Riggs, Dancu & Rothbaum, 1993), as did 88% of an American sample of abused women seeking emergency restraining orders. Rothbaum, Foa, Riggs, Murdock, and Walsh (1992) found that two-thirds of women met PTSD criteria one month post-incident, with referrals coming from a variety of sources, including mental health professionals.

When outlier findings appear, there is the possibility that extraneous or biasing factors have operated. Daly and Johnston (2002) reported that 67% of a sample of individuals held hostage in an Irish pub for three hours had chronic PTSD. Referral bias probably contributed to these results because all assessed individuals had been sent by plaintiff counsel for forensic assessments. Blanchard and Hickling (1997) reported that 44.3% of their sample met PTSD criteria after car accidents. This figure can be compared with those derived from epidemiological studies reporting conditional risks of PTSD after serious accidents (e.g., 6.3% for males and 8.8% for females; Kessler, Sonnega, Bromet, Hughes & Nelson, 1995). Once again, sampling issues probably contributed to high rates of PTSD. Studies by Blanchard and colleagues were conducted in a specialty clinic for motor vehicle accident victims, many of whom were referred by attorneys. Separate research from the same clinic demonstrated that litigation status was one of four factors most strongly predicting a PTSD diagnosis (Blanchard et al., 1996). Further, the ability of actors to feign PTSD when assessed in the clinic (Hickling, Blanchard, Mundy & Galovski, 2002) raises the specter of inflated prevalence rates associated with malingering (Rosen, 2006), when litigation is involved.

Although the majority of individuals who experience high-magnitude adverse events may not meet PTSD symptom criteria, the dose–response model has some utility. Overall, PTSD results more frequently after Criterion A events than it does after events that fail to meet the criterion’s definition (Kilpatrick et al., 1998). A classic illustration of a dose effect was obtained in an epidemiological survey of Manhattan residents after the terrorist attacks of September 11, 2001 (Galea et al., 2002). Approximately two months post-incident, 20% of surveyed individuals living South of Canal Street (in close proximity to the World Trade Center) reported reactions that met PTSD symptom criteria, compared with 6.8% of surveyed residents living North of Canal Street. Dohrenwend et al. (2006) recently documented a strong dose–response relationship in a reanalysis of NVDRS data. There also is the generally established finding that direct experiences of trauma lead to higher rates of disorder than vicariously experienced events (e.g., Breslau et al., 1998).

Despite such evidence consistent with the dose–response model, the vast literature on risk factors does not consistently support a monotonic relationship between stressor magnitude and clinical status, nor the assumption that Criterion A events are the “uniformly most potent” contributor to outcome (e.g., Bowman, 1997; Bowman & Yehuda, 2004; Brewin et al., 2000; Ozer et al., 2003). Consider, for example, how “trauma severity” and “stressor intensity” are relevant to the dose–response and adversity–stress models. Brewin et al. (2000) conducted a meta-analysis of risk factors for PTSD that included the variable of trauma severity. Overall, effect sizes for all risk factors were modest, whereas factors associated with the event including trauma severity exerted somewhat stronger effects. One limiting consideration, however, is that 77% of studies used by Brewin et al. in their meta-analysis relied on retrospective reporting. A problem with retrospective reporting is that PTSD-diagnosed subjects tend to demonstrate recall bias, such that they remember events as more traumatic over time (King et al., 2000; Roemer, Litz, Orsillo, Ehlich & Friedman, 1998; Schwarz, Kowalski & McNally, 1993; Southwick, Morgan, Nicolaou & Charney, 1997). The issue of recall bias confounds the question of interest, and may produce an artificial relationship between recalled severity of an event and clinical outcome.

Among prospective studies, methodological concerns still apply. Most important, the majority of prospective studies define trauma severity to include extent of physical injury. Defining trauma severity in this manner confounds event characteristics with event consequences, which is particularly troubling when extent of injury itself tends to be associated with PTSD status (March, 1993). Even with this confound, one that should produce a bias toward supporting a dose–response relationship, prospective studies on trauma severity have yielded mixed results. Some studies report a statistically significant relationship with clinical outcome (e.g., Epstein, Fullerton & Ursano, 1998; Frommberger et al., 1998; Koopman, Classen & Spiegel, 1994; Mayou, Ehlers & Bryant, 2002), whereas others do not (e.g., Ehlers, Mayou & Bryant, 1998; Perry, Difede, Musngi, Frances & Jacobsberg, 1992; Shalev, Peri, Canetti & Schreiber, 1996; Ursano et al., 1999). In some instances, studies failing to find a significant relationship still yield positive, albeit medium to small, effect sizes (e.g., calculated Cohen’s $d=.57$ in Perry et al., 1992; .27 in Shalev et al., 1996). Nevertheless, the
general conclusion remains that research provides at best inconsistent support for a dose–response relationship between clinical outcome and event/injury severity.

When trauma severity is found to contribute to clinical outcome, regardless of how that variable is defined, it still is not the greatest or “uniformly most potent” (Meehl, 1977) contributor. This conclusion holds across studies, even though specific findings may differ. For example, Brewin et al. (2000) found social support to be the strongest predictor of PTSD status (weighted \(r=.40\)) compared with trauma severity (weighted \(r=.23\)). Ozer et al. (2003) found peritraumatic dissociation (dissociation occurring during the trauma itself) to be the strongest predictor (weighted \(r=.35\)), as compared with perceived life threat (weighted \(r=.26\)). In both meta-analyses, the combined influence of prior trauma history, prior adjustment, family history of psychopathology, strength of peritraumatic reactions, and social support contributed substantially more to outcome than specific event characteristics.

Evidently, Criterion A events are neither necessary nor sufficient to produce the defined clinical syndrome of PTSD, nor are they uniformly most potent in their influence. This conclusion holds true for the objective (A1) and subjective (A2) components of a traumatic stressor. It is therefore surprising that a major report on PTSD issued by the Institute of Medicine (IOM, 2006) gave no consideration to the multiple and complex issues surrounding Criterion A. Instead, the IOM report merely concluded that, “the necessary cause of PTSD is by definition a traumatic event” (p. 23). Yet as we have seen, the presumed link between Criterion A and subsequent signs and symptoms is critical to the question of PTSD’s construct validity.

In retrospect, the often limited influence of Criterion A events should not be entirely surprising. As Turkheimer and Waldron (2000) observed in their meta-analysis of nonshared environmental influences (i.e., variables that tend to decrease the similarity among family members), the contributions of specific nonshared environmental events (e.g., stressors) tend to be quite low in magnitude (see also Meehl, 1978, for a discussion of the “random walk”). Most traumatic stressors, if not all, appear to be broadly consistent with Turkheimer and Waldron’s conclusion that nonshared environmental influences rarely exert large or uniformly potent main effects on emotional adjustment or personality.

1.3. Changing definitions, A2, and criterion creep

The original definition of Criterion A in the DSM-III (APA, 1980) was a single sentence that referred to the “Existence of a recognizable stressor that would evoke significant symptoms of distress in almost everyone” (p. 238). By the time of DSM-IV’s publication (APA, 1994), substantial extensions of Criterion A had occurred:

The person has been exposed to a traumatic event in which both of the following were present: (1) the person experienced, witnessed, or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others. (2) the person’s response involved intense fear, helplessness, or horror. (p. 467).

Text from the DSM-IV clarified that “confronted” events did not need to be directly experienced or even witnessed. Instead, an individual could merely learn about an event that included “violent personal assault, serious accident, or serious injury experienced by a family member or a close friend.” (p. 464) Thus, in the span of three DSM editions, the universe of potentially traumatic events went from direct experiences to only having to hear about a severe misfortune befalling others. Breslau and Kessler (2001) found that the expanded definition of Criterion A in the DSM-IV increased by 59% the total number of events considered traumatic.

Rosen (2004) referred to the expansion of events subsumed under Criterion A as “criterion creep,” while McNally (2003a) used the term “conceptual bracket creep.” With the expanded definition, an individual who sees horrific events on the news can technically develop PTSD (e.g., Ahern, Galea, Resnick & Vlahov, 2004; Eth, 2002; Pfefferbaum, Pfefferbaum, North & Neas, 2002; Propper, Stickgold, Keeley & Christman, 2007). A recent study reported on the development of delayed PTSD associated with watching television one year after the September 11, 2001 attacks (Bernstein et al., 2007). Zimering, Gulliver, Knight, Munroe, and Keane (2006) found that 4.6% of mental health relief workers met symptomatic criteria for PTSD after hearing narratives from 9/11 survivors, without any personal connection to victims. These findings relate to an expanding literature on “vicarious traumatization” (Sabin-Farrell & Turpin, 2003).

In addition to DSM’s “official” expansion of events subsumed under Criterion A, various unofficial proposals have been advanced. For example, Avina and O’Donohue (2002) argued that typically non-traumatic instances of sexual
harassment such as crude jokes in the workplace could be considered traumatic because victims might worry, “since he is doing this to me, what else is he capable of?” (p. 72). Rosen (2004) maintained that Avina and O’Donohue effectively extended the range of events subsumed under Criterion A to the realm of expectations, a type of criterion creep that created the conceptual equivalent of “pre-traumatic” stress disorder.

Pre-traumatic stress disorder relates to the second component (A2) of Criterion A: the subjective experience of fear, helplessness, or horror. In this context, the arguments advanced by Avina and O’Donohue (2002) can be viewed as extending the significance of subjective reactions to the perception of future risk. The inclusion of a subjective component to define Criterion A has received empirical support (e.g., Brewin, Andrews & Rose, 2002), in part because the extent of subjective emotion to trauma is strongly associated with other aspects of event magnitude (Weathers & Keane, 2007). Also, Criterion A2 appears particularly useful in predicting posttraumatic morbidity in so far as the absence of a strong emotional component is a strong indicator that PTSD will not develop. In other words, A2 has high negative predictive value (Breslau & Kessler, 2001; Creamer, McFarlane & Burgess, 2005; Schnurr, Spiro, Vielhauer, Findler & Hamblen, 2002). It also has been found that negative emotions other than fear, helplessness, and horror (e.g., anger, shame) are associated with the development of PTSD (Brewin et al., 2000). Such findings suggest additional avenues for the expansion of A2’s defining features.

The potential for criterion creep associated with Criterion A2 is further illustrated in a study by McNally et al. (2004). These researchers found that people who believed they had been abducted by space aliens were reactive physiologically to “trauma” script-driven imagery. McNally and colleagues did not argue that these findings supported the reality of alien abductions. Rather, they observed that merely believing that one has been traumatized can generate emotional responses much like those exhibited by true trauma victims (e.g., veterans who experienced intense combat). These findings raise the possibility that PTSD can result from any event, real or imagined, if the affected individual believes it was sufficiently stressful.

The multiple and fundamental issues raised in various attempts to define trauma were recently termed by Weathers and Keane (2007) the “Criterion A problem.” One approach to dealing with this problem would be to create the greatest possible expansion of Criterion A and include all stressful events as potential contributors to the PTSD clinical syndrome. This approach is consistent with current findings and is not without precedent. In preparation for the DSM-IV, various proposals for redefining Criterion A were entertained, including the possibility of dropping the concept of a distinct class of stressors (Kilpatrick et al., 1998). Specifically, the group included a “nonrestrictive option,” for which “No a priori exclusion of the types of events capable of leading to PTSD” would be required (p. 839). This nonrestrictive option was rejected in the DSM-IV, but continues to be recommended by some (Maier, 2006). Should future editions of the DSM adopt a nonrestrictive option and eliminate Criterion A, then PTSD as originally conceived would cease to exist. If that happens, it logically follows that the rationale for PTSD would fall to the defining set of symptom criteria.

2. Is PTSD a distinct clinical syndrome?

Maser and Cloninger (1990) observed that no individual sign or symptom in psychiatry is two-way pathognomonic – that is, both necessary and sufficient – for a specific psychiatric disorder. Instead, diagnoses within the DSM are based on multiple nonspecific signs and symptoms. Nevertheless, to stand as a distinct clinical syndrome, a diagnosis must demonstrate that it is not simply an amalgam of symptoms and processes associated with already extant disorders (Robins & Guze, 1970; March, 1990).

Since PTSD was introduced in the DSM-III, the symptom criteria that define the construct have been grouped into three clusters (Criteria B through D). Table 1 presents the original 12 symptom criteria contained in the DSM-III (APA, 1980) and the 17 symptom criteria presented in the DSM-III-R (APA, 1987) to illustrate significant differences across the two editions. With DSM-IV (APA, 1994), the 17 DSM-III-R symptoms remained, although descriptors were edited and criterion D6 moved to become B5. The three clusters as they now stand reflect reexperiencing symptoms (Criterion B), efforts at avoidance and emotional numbing (Criterion C), and reactions associated with hyperarousal (Criterion D). In each instance, the symptom clusters are assumed to reflect underlying processes of anxiety, with PTSD being listed in the DSM among the Anxiety Disorders (Barlow, 2002; Jones & Barlow, 1990).

Over the years, numerous factor analyses have examined the structure of PTSD symptoms. Some of these studies have confirmed three symptom dimensions (Anthony, Lonigan & Hecht, 1999; Cordova, Studts, Hann, Jacobsen & Andrykowski, 2000; Foa, Riggs & Gershuny, 1995; Ventureyra, Yao, Cottraux, Note & Mey-Guillard, 2002). Other studies have found as few as two (Buckley, Blanchard & Hickling, 1998; Taylor, Kuch, Koch, Crockett & Passey,
Table 1
Symptom criteria for PTSD in the 1980 and 1987 editions of the DSM

<table>
<thead>
<tr>
<th>DSM-III</th>
<th>DSM-III-R</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B. Reexperiencing of the trauma as evidenced by at least one of the following:</strong></td>
<td><strong>B. The traumatic event is persistently reexperienced in at least one of the following ways:</strong></td>
</tr>
<tr>
<td>(1) recurrent and intrusive recollections of the event</td>
<td>(1) recurrent and intrusive distressing recollections of the event</td>
</tr>
<tr>
<td>(2) recurrent dreams of the event</td>
<td>(2) recurrent distressing dreams of the event</td>
</tr>
<tr>
<td>(3) sudden acting or feeling as if the traumatic event were reoccurring, because of an association with an environmental or ideational stimulus</td>
<td>(3) sudden acting or feeling as if the traumatic event were recurring</td>
</tr>
<tr>
<td><strong>C. Numbing of responsiveness to or reduced involvement with the external world, beginning some time after the trauma, as shown by at least one of the following:</strong></td>
<td><strong>C. Persistent avoidance of stimuli associated with the trauma or numbing of general responsiveness (not present before the trauma), as indicated by at least three of the following:</strong></td>
</tr>
<tr>
<td>(1) markedly diminished interest in one or more significant activities</td>
<td>(1) efforts to avoid thoughts or feelings associated with the trauma</td>
</tr>
<tr>
<td>(2) feeling of detachment or estrangement from others</td>
<td>(2) efforts to avoid activities or situations that arouse recollections of the trauma</td>
</tr>
<tr>
<td>(3) constricted affect</td>
<td>(3) inability to recall an important aspect of the trauma (psychogenic amnesia)</td>
</tr>
<tr>
<td><strong>D. At least two of the following symptoms that were not present before the trauma:</strong></td>
<td><strong>D. Persistent symptoms of increased arousal (not present before the trauma), as indicated by at least two of the following:</strong></td>
</tr>
<tr>
<td>(1) hyperalertness or exaggerated startle response</td>
<td>(1) difficulty falling or staying asleep</td>
</tr>
<tr>
<td>(2) sleep disturbance</td>
<td>(2) irritability or outbursts of anger</td>
</tr>
<tr>
<td>(3) guilt about surviving when others have not, or about behavior required for survival</td>
<td>(3) difficulty concentrating</td>
</tr>
<tr>
<td>(4) memory impairment or trouble concentrating</td>
<td>(4) hypervigilance</td>
</tr>
<tr>
<td>(5) avoidance of activities that arouse recollection of the traumatic event</td>
<td>(5) exaggerated startle response</td>
</tr>
<tr>
<td>(6) intensification of symptoms by exposure to events that symbolize or resemble the traumatic event</td>
<td>(6) physiologic reactivity upon exposure to events that symbolize or resemble an aspect of the traumatic event</td>
</tr>
</tbody>
</table>

1998) and as many as five (Watson et al., 1991) factors. The majority of factor analytic studies have generated best fit models yielding four factors that are often, but not always, labeled as reexperiencing, effortful avoidance, emotional numbing, and hyperarousal (Amdur & Liberzon, 2001; Asmundson et al., 2000; Asmundson, Stapleton & Taylor, 2004; Baschnagel, O’Connor, Colder & Hawk, 2005; DuHamel et al., 2004; King, Leskin, King & Weathers, 1998; McWilliams, Cox & Asmundson, 2005; Simms, Watson & Dobbeling, 2002).

Factor analytic studies and the models they generate are of considerable interest for ascertaining the internal structure of PTSD as a syndrome, but they do not determine if the diagnosis of PTSD best carves nature at its joints. Rather, these studies speak to relationships and covariations among a selected set of responses that humans experience to adverse events.

### 2.1. Variability of symptom patterns

A difficulty in establishing that PTSD is a distinct syndrome is the variability that occurs in clinical presentations among PTSD-diagnosed patients. Foa et al. (1995) observed how symptom criteria for PTSD render the diagnosis heterogeneous such that two individuals can receive the diagnosis yet not share a single symptom. This is possible as a result of there being 17 symptom criteria, only 6 of which are required for the diagnosis. Thus, one individual can report nightmares, loss of interest in activities, emotional numbing, sense of a foreshortened future, sleep disturbance and concentration problems. A second individual can report intrusive thoughts, avoidance of reminders and discussions of the traumatic event, inability to recall important details of the event, irritability, and hypervigilance.
Marked variability in symptom presentations does not by itself demonstrate that PTSD is not a distinct diagnostic entity. Variable presentations may all be alternative manifestations or formes frustes of a shared underlying etiology. Indeed, well-established disorders in medicine illustrate the point that individuals with the same disease can have varied presentations. One gastrointestinal problem, celiac disease, exhibits so many symptom profiles it has been called the “great imposter” (Lee & Green, 2006). When symptom patterns are not pathognomonic (e.g., they do not definitively indicate either the presence or absence of a distinct disorder) then a disease entity must be validated in some other manner. In the case of celiac disease, a specific etiology was determined (food intolerance to the gluten protein contained in certain grains) and this causal mechanism was connected with a distinct pathophysiology (damage to the lining of the small intestine). In sharp contrast to the history of celiac disease, attempts to identify a specific etiology for PTSD (Criterion A) have not received empirical support. Also meeting with failure have been numerous attempts to identify a distinct pathophysiology or other validating marker of the PTSD construct, an issue discussed later in the paper. Without a specific etiology, and without a distinct pathophysiology, the marked variability in symptom presentation observed among PTSD diagnosed patients challenges the construct.

### 2.2. Comorbidity

Yehuda and McFarlane (1995) observed that studies of comorbidity call into question whether PTSD can be clearly differentiated from other psychiatric disorders. This can be a problem because extremely high levels of comorbidity may challenge a condition’s syndromal independence and construct validity (Lilienfeld, Waldman & Israel, 1994). As observed by the IOM (2006), “Comorbidity, by itself, does not preclude the validation of PTSD as a distinct disorder, but it makes the process of demonstrating its distinctiveness more difficult.” (pp. 50–51). Moreover, extensive comorbidity raises the largely neglected question of whether PTSD possesses incremental validity (Sechrest, 1963) above and beyond co-occurring diagnoses for clinically and theoretically meaningful external criterion variables.

Young (2004a) referred to the high comorbidity of PTSD with other disorders as the “soft under-belly of this diagnostic classification.” (p. 388) The problem here is not simply that many PTSD patients carry additional diagnoses such as major depressive disorder (MDD), panic disorder (PD), generalized anxiety disorder (GAD), alcoholism (Deering, Glover, Ready, Eddleman & Alarcon, 1996; Keane & Wolfe, 1990; Kessler et al., 1995) and obsessive–compulsive disorder (Helzer et al., 1987). Rather, the problem is that many symptom criteria that define PTSD also serve to define the very disorders with which PTSD most frequently co-occurs (Rosen, Spitzer & McHugh, 2008; Spitzer et al., 2007). In this context, King and King (1991) observed: “The long-term survival of the construct [PTSD] certainly depends on its ability to stand alone, and hence the concern for discriminant validity” (p. 117). The Kings emphasized the critical task of distinguishing PTSD from other diagnostic constructs to “thereby confirm its place as a unique diagnostic entity.” (p. 121).

Consider how the DSM-IV defines symptom criteria B-4 for PTSD as “intense psychological distress at exposure to internal or external cues that symbolize or resemble an aspect of the traumatic event,” while it defines symptom criteria A for Specific Phobia as “marked and persistent fear that is excessive or unreasonable, cued by the presence or anticipation of a specific object or situation.” Symptom criteria C-4 for PTSD is defined as “markedly diminished interest or participation in significant activities,” while symptom 2-A for major depression is defined as “markedly diminished interest or pleasure in all, or almost all activities most of the day, nearly every day.” One of the hallmark symptoms of PTSD concerns the problem of intrusive thoughts (criterion B-1). Yet studies have found that on both quantitative and qualitative measures there are few differences between the intrusive memories of depressed and PTSD patients (Brewin, 1998; Reynolds & Brewin, 1999). Mellman and Davis (1985) observed that the phenomenology of PTSD flashbacks (criteria B-2) meet the DSM-III criteria for panic attacks. Bernstein (2001) found that highly vivid involuntary memories (flashbacks) are not limited to trauma or even to emotionally negative events.

No publication has addressed the issue of whether the dual diagnoses of specific phobia and depression can account for at least a portion of PTSD-diagnosed cases. McNally and Saigh (1993) discussed several case reports of individuals who had developed circumscribed phobic reactions without fulfilling the symptom criteria for PTSD. McNally and Saigh discussed how specific phobias were thereby distinguishable from PTSD. Yet, these case examples do not clarify how best to diagnose individuals who are broadly affected by a traumatic event and who present with a range of reactions that meet criteria for a specific phobia and depression (or alternatively, the diagnosis of PTSD).

Keane, Wolfe, and Taylor (1987) noted that Goodwin and Guze (1984) questioned the validity of PTSD because of its substantial phenomenological overlap with major depression and GAD. At the time of their 1987 paper, Keane et al.
reported that two studies were in progress. They concluded, “Clearly, much information needs to be obtained before we can conclude that PTSD can be delimited from other homogeneous psychiatric groups” (p. 35). One of the studies referenced by Keane et al. was eventually published (Keane, Taylor & Penk, 1997). In this study an attempt was made to differentiate PTSD from major depression and GAD by asking clinicians to rate the extent to which 90 listed symptoms were associated with the competing disorders. It was found that clinicians could distinguish PTSD symptoms from those of MDD and GAD, which by implication was taken to support the validity of the PTSD construct. This finding was never tested on actual patients.

In reality, the Keane et al. study demonstrated only that clinicians can correctly apply accepted convention. This outcome should not be unexpected, particularly when the only event-specific diagnosis (i.e., PTSD) was associated with event specific symptom descriptors (e.g., “Suddenly acting or feeling as if an event were recurring;” “Recurring dreams of an event that happened”). Event-specific content may define PTSD and guide clinicians who reliably apply the diagnosis, but it provides at best weak validation for the construct. Put more generally, and applying an old adage, reliability does not necessarily imply validity.

The study by Keane et al. (1997), accompanied as it was with significant limitations, was the only one considered by the Institute of Medicine in their assessment of whether PTSD could be differentiated from other disorders (IOM, 2006). The IOM report observed that Keane et al., “did not demonstrate that health professionals can distinguish PTSD, GAD, and MDD from each other in patients, but [they] did show that features of PTSD are conceptually distinct from the other disorders.” On this basis alone, members of the IOM committee concluded that progress had been made in demonstrating that PTSD is distinct from disorders with similar symptoms.

The IOM report ignored several studies that suggest PTSD and depression arise from similar predictive variables and a shared vulnerability (Breslau, Davis, Peterson & Schultz, 2000; O’Donnell, Creamer & Pattison, 2004), such that these disorders should not be viewed as entirely distinct conditions. O’Donnell et al. proposed a general stress factor to account for these states with resulting distress determined by event characteristics, individual characteristics, negative cognitive appraisals, and the level of initial reactions (symptoms over the first week posttrauma). This framework is consistent with studies on the latent structure of anxiety and depression (Brown, Chorpita & Barlow, 1998a; Clark & Watson, 1991; Krueger, 1999) and a dimensional model (Forbes, Haslam, Williams & Creamer, 2005; Ruscio, Ruscio & Keane, 2002) that conceptualizes PTSD as the upper end of a stress–response continuum rather than a taxon or discrete clinical syndrome (Broman-Fulks et al., 2007; Forbes et al., 2005; Ruscio et al., 2002).

Indeed, research using taxometric statistical methods (Meehl & Golden, 1982) suggests that posttraumatic stress symptoms fall on a continuum of severity that includes everyday symptoms of stress at one end, and more severe PTSD symptomatic criteria, at the other. Further, one is left with a multitude of reactions that are themselves associated with previously known processes (e.g., depression, anxiety) and co-occur in a variety of permutations and patterns (Bogenschutz & Nurnberg, 2000; Brown, 2001). Such findings raise the thorny issue of whether PTSD should be rejoined with the disorders with which it shares common features (e.g., specific phobia, generalized anxiety disorder, panic disorder, major depression) or whether it should remain apart.

2.3. Emotional numbing

Despite symptom overlap, the possibility remains that PTSD and depression are independent responses to trauma (Blanchard, Buckley, Hickling & Taylor, 1998; Erickson, Wolfe, King, King & Sharkansky, 2001; Franklin & Zimmerman, 2001). Also, studies have suggested that emotional numbing (EN) may be relatively specific to PTSD (Breslau, Reboussin, Anthony & Storr, 2005; Kashdan, Elhai & Frueh, 2006; Ruscio, Weathers, King & King, 2002). As previously reviewed, factor analytic studies most often generate a best fit model containing four factors in which avoidance and EN are no longer combined into symptom cluster C (e.g., Asmundson et al., 2004). Support for separating avoidance and EN as separate processes also comes from animal studies and the models they generate, if one views the analgesia observed in stressed animals as paralleling the numbing symptoms of PTSD (Foa, Zinbarg & Rothbaum, 1992). Further, research on PTSD-diagnosed individuals has shown that EN (a) may be distinct from depression (Flack, Litz, Hseih, Kaloupek & Keane, 2000; Litz et al., 1997); (b) contributes to the prediction of PTSD after depression has been statistically controlled (Feeny, Zoellner, Fitzgibbons & Foa, 2000); and (c) may relate to PTSD severity and low cortisol levels (Mason et al., 2001). These findings support the observation that EN-related symptoms may be particularly important to the etiology of PTSD (Breslau et al., 2005; Foa et al., 1995; Litz, 1992; Litz & Gray, 2002). In this context, Foa et al. (1995) proposed that individuals who experience trauma but do not evidence
EN should be diagnosed with another anxiety condition rather than PTSD. This proposal, which could significantly reduce the heterogeneity among PTSD diagnosed samples, requires empirical corroboration.

Findings on EN as a PTSD symptom factor distinct from depression are somewhat puzzling because of obvious overlap in criteria that define the two constructs. EN among PTSD patients is defined by symptom criteria C4–6, which include diminished interest in activities, detachment or estrangement from others, and restricted range of affect. The similarity of these features to depression is evident in a review of criteria from DSM-IV or in a review of items on such depression measures as the Beck Depression Inventory. Because a core feature of EN appears to be the inability to experience positive emotions, we recommend further research to ascertain the similarities and differences between EN and anhedonia (Kashdan et al., 2006; Litz & Gray, 2002). Nevertheless, the proposal that EN is an emotional process distinct from depression has received some support and suggests the possibility of a more restrictive definition of PTSD (Feeny et al., 2000; Foa et al., 1995).

2.4. The matter of bereavement

The issue of overlapping symptoms extends further when one considers the phenomenon of bereavement. Bereavement encompasses many of the same reactions (e.g., intrusive thoughts, distress at reminders, anxiety) that constitute the symptom criteria for PTSD, although the content and focus of recurring thoughts and distressed emotions can differ (Burnett et al., 1994; Raphael, Martinek & Wooding, 2004; Stroebe, Hansson, Stroebe & Schut, 2001). The DSM-IV (APA, 2000) lists bereavement as a condition other than a mental disorder that may be the focus of clinical attention. Of greater concern clinically are “complicated bereavement” and “traumatic grief” (e.g., Jacobs, 1999; Raphael et al. 2004), although these are not formally recognized in the DSM. Rather, the DSM-IV applies extant diagnoses to account for these more disturbed states. For example, text in the DSM-IV-TR (APA, 2000) notes that duration and expression of “normal” bereavement can vary among different cultural groups. The DSM then clarifies that in the context of North American culture a diagnosis of major depressive disorder is generally not provided after the loss of a loved one, even if the bereaved individual’s reactions meet criteria for this disorder, unless the reactions are present for at least two months. No empirical basis is provided for this suggested time frame.

Several studies have examined the occurrence of PTSD symptom criteria among bereaved individuals, finding that PTSD-related reactions are most likely when the death of a loved one results from violent (e.g., homicide) as opposed to natural (e.g., death from illness) causes (Barry, Kasl & Priegerson, 2002; Bonanno & Kaltman, 1999; Kaltman & Bonanno, 2003; Momartin, Silove, Manicavasagar & Steel, 2004; Zisook et al., 1998). Yet such findings do not demonstrate that PTSD-related symptoms involve an underlying process distinct from bereavement. They show only that certain conditions or circumstances associated with death are likely to be associated with characteristic reactions. In this context, it should be recognized that after the homicide of a loved one, bereaved individuals face extensive situational stressors (e.g., police reports, court trials) that differ from the circumstances of a natural death (e.g., heart attack). These concerns and others may contribute to differences in reported response patterns.

The overlapping reactions of bereavement and PTSD deserve more attention than has generally been accorded by researchers. If bereavement and the parallel concept of grief are extended to include a broader domain of loss, such as loss of function or loss of one’s sense of the world (Janoff-Bulman, 1992), then these concepts become even more relevant to an analysis of posttraumatic reactions.

3. Attempts to identify distinctive markers

Researchers have tried to validate the PTSD construct by identifying one or more distinguishing markers (e.g., laboratory or biological indicators) that might serve as the disorder’s “signature” and thereby validate the construct as separable from overlapping conditions. In this way, it was hoped that PTSD could be firmly established as a true disease in nature (Young, 1995, 2001, 2004b). The search for distinctive features or signatures that distinguish one psychiatric construct from another is hardly unique to PTSD. Nevertheless, such endeavors have assumed particular urgency in research on PTSD, as other attempts to validate the construct (e.g., identifying a specific etiology or a defining pathognomonic symptom) have failed to achieve adequate levels of empirical support. In this context, we are reminded of D.W. King’s and L.A. King’s (1991) concern for the survival of PTSD and their urging “researchers to implement a systematic and deliberate program of discriminant validity research” (p. 121).

The next several sections of this manuscript provide an overview of efforts to identify a distinctive marker for PTSD.
3.1. Neuroendocrine findings

Meehl (1977) noted that the majority of disease entities in the history of medicine approximate an “ideal type” (p. 34) that is defined by a conjunction of its pathology and etiology. In contrast to many disorders in organic medicine, few if any mental disorders are defined in this fashion (Lilienfeld et al., 1994). Indeed, “hard” physiological data that could establish the reality of a mental disorder are notoriously hard to come by. It was therefore with considerable interest that the “cortisol hypothesis” was advanced to differentiate PTSD from other disorders.

It had long been assumed in the field of general stress studies that adverse events lead to activation of the hypothalamic–pituitary–adrenal (HPA) axis, as reflected in elevated levels of cortisol among stressed subjects. Contrary to that long held model, several studies documented significantly lower urinary cortisol among groups of PTSD diagnosed patients (Mason, Giller, Kosten, Ostroff & Podd, 1986; Yehuda, Southwick, Nussbaum, Giller & Mason, 1990). These findings contrasted with higher cortisol levels (hypercortisolemia) obtained among patients with major depression (e.g., Gillespie & Nemeroff, 2005; Krishnan, 1993), thus providing support for the discriminant validity of PTSD from overlapping disorders. Findings of “hypocortisolism” among PTSD-diagnosed patients led to the view that a specific HPA axis dysfunction could come to define an underlying physiopathology for the disorder. Since the original studies, low cortisol levels have been reported not only for PTSD diagnosed patients (Yehuda et al., 1995) but for their offspring (Yehuda et al., 2000, 2007), a finding interpreted by Yehuda and colleagues to suggest a vulnerability marker related to parental PTSD. Boscarino (1996) provided another study reporting lower cortisol levels among veterans with current PTSD. However, this finding was not replicated for veterans with a lifetime PTSD diagnosis, thereby raising several possibilities including the reversal of altered neuroendocrine function with remission.

The finding that groups of PTSD patients exhibit mean cortisol levels lower than comparison subjects has required revision of long-standing stress models (e.g., Selye, 1946). Research also has led to a better appreciation for the complexity of stress reactions (Rasmussen, Vythilingham & Morgan, 2003; Yehuda, 2002). At the same time, multiple considerations limit the usefulness of cortisol levels as a marker that might validate the PTSD construct. For example, not all researchers have found lower cortisol levels among groups of PTSD patients (Liberzon, Abelson, Flagel, Raz & Young, 1999; Pitman & Orr, 1990; Yehuda, Lowy, Southwick, Shaffer & Giller, 1991). Indeed, some have reported the opposite. Liberzon et al. (1999) found significantly higher baseline cortisol among PTSD diagnosed patients than among normals (p=.02; calculated Cohen’s d=.82). Pitman and Orr (1990) found that baseline 24-hour urinary-free cortisol was significantly higher in PTSD than in healthy subjects (p=.02; calculated Cohen’s d=.84). Yehuda et al. (1991) reported non-significant differences between PTSD subjects and controls, without indicating the direction of effects or providing the mean or standard deviation for controls.

Epidemiologic studies have failed to document lower cortisol levels among individuals who report exposure to trauma and who meet the symptom criteria for PTSD (Young & Breslau, 2004a, 2004b; Young, Tolman, Witkoński & Kaplan, 2004). Breslau (2006) observed that the non-replication of clinical studies in community based samples is a particularly serious issue for the cortisol hypothesis of PTSD. As noted by Breslau, selection and referral biases can occur with clinic based samples, where comorbidity tends to be especially frequent (Berkson, 1946; Roberts, Spitzer, Delmon & Sackett, 1978). Therefore, although the first clues for an association between a variable of interest and a disease often come from clinical samples, confirmation that the association is valid must come from other sources such as unselected community samples (also see Goodman, Lahey, Fielding, Dulcan & Regier, 1997).

Variability of cortisol findings within and among patients (Mason et al., 2002) appears to be the norm. Even when mean cortisol levels for PTSD-diagnosed groups differ from controls, the findings can reflect group mean scores and fictive patients (i.e., prototypical patients who rarely exist in nature) rather than the majority of actual individuals. Young (2004a) provided an example of this concern by analyzing data reported by Yehuda et al. (1990). In that study, Yehuda et al. found significantly lower urinary cortisol levels in 16 non-medicated male combat veterans diagnosed with PTSD compared with 16 age-comparable nonpsychiatric male control subjects (p<b.001; calculated Cohen’s d=1.22). Young observed that only 2 of the 16 PTSD patients in Yehuda et al.’s study had cortisol levels that fell below the lowest score obtained in the comparison group. Put another way, although group means differed, the independent variable had little diagnostic utility.

Further concerns regarding the implications of cortisol findings with PTSD patients are raised by studies that report low cortisol levels among adults with fibromyalgia (Fries, Hesse, Hellhammer & Hellhammer, 2005; Griep et al., 1998) and chronic fatigue syndrome (Demitrak et al., 1991; Roberts, Wessely, Chalder, Papadopoulos & Cleave, 2004), and children with conduct disorders (Oosterlaan, Geurts, Knowl & Sergeant, 2005; Shoal, Giancola & Kirillova, 2003). It is
unclear how a physiological finding that occurs in diverse patient groups, possesses minimal diagnostic utility, and applies to a small minority of group members can serve to establish PTSD as a distinct disorder.

### 3.2. Neuroanatomy

In an article, “Does Stress Damage the Brain?,” Bremner (1999) reviewed animal studies that revealed a relationship between stress and reduced volumes of the hippocampus, an area of the brain associated with learning and memory. Bremner argued that this ostensible damage was related to high levels of glucocorticoids and possibly other neurotransmitters. Bremner also reviewed studies that found reduced hippocampal volumes among combat veterans (Bremner et al., 1995; Gurvits et al., 1996) and victims of childhood abuse (Bremner et al., 1997). Based on such findings, Bremner (1999) proposed that traumatic stressors can damage the brain, thereby leading to PTSD: “stress-induced brain damage underlies and is responsible for the development of a spectrum of trauma-related psychiatric disorders, making these psychiatric disorders, in effect, the result of neurological damage” (Bremner, 2002, p. 4).

Additional studies (Bremner et al., 2003; Gilbertson et al., 2002; Stein, Koverola, Hanna, Torchia & McClarty, 1997a; Villarreal et al., 2002) and a meta-analysis (Kitayama, Vaccarino, Kutner, Weiss & Bremner, 2005) have provided support for the finding of reduced hippocampal volumes among PTSD patients. Vermetten, Vythilingam, Southwick, Charney, and Bremner (2003) extended the possible implications of research on hippocampal volumes by reporting that treatment of 23 PTSD diagnosed patients with paroxetine (Paxil) was associated with improved verbal declarative memory and a 4.6% increase in mean hippocampal volume. There is precedent for finding in that reversals in hippocampal atrophy have been demonstrated after treatment of patients with Cushing’s disease (Starkman et al., 1999) and after long-term use of antidepressants (Malberg, 2004). At the same time multiple problems befall the study by Vermetten et al. (2003): they did not employ a comparison group; improvements in verbal declarative memory could have been an artifact of practice effects; there was no significant correlation between change in PTSD symptoms and change in hippocampal volume (r not reported by authors); and several (n = 6) of the 23 patients evidenced smaller hippocampal volumes on retesting (displayed in graph but not reported in text).

The report on PTSD by the Institute of Medicine, cited only the one study by Bremner et al., 1995), and observed that reduced hippocampal volume “has been found to be associated with PTSD.” (p. 54). Yet a challenge to Bremner’s (1999, 2002) claim arises when one considers alternative hypotheses. For example, Stein et al. (1997a,b) suggested that pre-incident reductions in hippocampal volume could predispose individuals to develop morbidity after trauma. At the time of Stein’s article, and for several years thereafter, no studies on premorbid hippocampal size addressed the question of cause and effect. Suggestive studies that bore on the issue (Pitman, 2001) had found significantly more pre-incident neurodevelopmental abnormalities, learning difficulties, and associated difficulties among PTSD abused women as compared with non-PTSD traumatized controls (Gurvits et al., 1993,2000). Studies also had found significantly lower performance on pre-trauma intelligence testing among PTSD veterans compared with non-PTSD controls (Macklin et al., 1998; McNally & Shin, 1995). These studies pointed to pre-incident neurological vulnerabilities that might adversely affect an individual’s ability to cope with traumatic events.

Gilbertson et al. (2002) used a case control design and examined male monozygotic twin pairs discordant for Vietnam combat exposure. In some twin pairs, the combat-exposed brother was diagnosed with PTSD, thereby providing for a number of comparisons and built-in controls for genetic effects. Gilbertson et al. replicated the finding of reduced hippocampal volume in PTSD subjects, as compared with non-PTSD trauma exposed subjects. However, they also found a pattern of similar volumes in matched twins (smaller volumes in brothers of PTSD subjects as compared with brothers of non-PTSD subjects). These findings support the hypothesis that reduced hippocampal volumes among PTSD patients represent a pre-incident vulnerability factor rather than damage caused by trauma.

Further tempering the claim that stress damages the brain are multiple failures to replicate findings of smaller hippocampal volume among PTSD diagnosed patients. (Bonne et al., 2001; DeBellis, Hall, Boring, Frustaci & Moritz, 2001; Fenmema-Notestine, Stein, Kennedy, Archibald & Jermigan, 2002; Golier, et al., 2005; Pederson et al., 2004; Schuff et al., 2001). An interpretation of research on neuroanatomy and PTSD also is confounded by studies that report reduced hippocampal volumes among patients diagnosed with depression (e.g., Bremner et al., 2000; Mervahaal et al., 2000; Sheline, Sanghavi, Mintun & Gado, 1999; Steffens et al., 2000), a condition that co-occurs substantially with PTSD. Other issues weaken the empirical basis for claims that stress damages the brain (Jelicic & Merckelbach, 2004). These issues include crude measurement methods that yield control values out of line with other investigations (Colchester et al., 2001; Kopelman, 2002), the confounding effects of comorbid alcoholism (Woodward et al., 2006a),
and the occasional inclusion of subjects who experienced loss of consciousness and the risk that short periods of hypoxia can result in hippocampal damage (Warden, Redier-Groswasser, Grafman & Salazar, 1995).

Finally, post-incident reductions in hippocampal volumes cannot logically account for the clinical syndrome of PTSD when posttraumatic reactions display rapid onset after the presumed etiologic event (e.g., Rothbaum et al., 1992; Shalev, 2002). North (2001) assessed individuals after the Oklahoma City bombing and found that 34% of the sample met criteria for PTSD. Among PTSD diagnosed individuals, onset of symptoms was rapid, with 76% of individuals reporting that symptoms arose the first day, and another 18% reporting that symptoms arose within the first week. If hippocampal atrophy is a gradual process (absent direct physical trauma) then stress-induced brain damage (Bremner, 2002) could not be responsible for posttraumatic symptoms that typically develop in the immediate aftermath of trauma. Further, the notion that high levels of early posttraumatic symptoms might result in progressive hippocampal damage, which then contributes to a failure to recover, is not consistent with the robust finding observed in the DSM (APA, 2000, p. 466) that 50% of PTSD-diagnosed individuals improve within three months of symptom onset (e.g., Rothbaum et al., 1992). Further, Bonne et al. (2001) employed a longitudinal design with brain scans performed on 37 individuals within a week of their traumatic event and again six months later. No change was observed within this time frame for right or left hippocampal volumes among ten participants who met PTSD criteria (e.g. right hippocampal volume at 1 week, x=3.95, sd=.42; right hippocampal volume at 6 months, x=3.98, sd=.42).

More recently, studies have focused on reduced volumes in other areas of the brain, such as the anterior cingulate (Kitayama, Quinn & Bremner, 2006; Woodward et al., 2006b; Yamasue et al., 2003). These findings raise the same concerns as those of research on the hippocampus. In this context, even the report on PTSD by the Institute of Medicine (IOM, 2006), concluded that, “No biomarkers are clinically useful or specific in diagnosing PTSD, assessing the risk of developing it, or charting its progression” (p. 46).

### 3.3. Neurocircuitry

An emerging area of research that may hold promise involves neuroimaging studies with PTSD-diagnosed individuals and patients with other anxiety disorders (e.g., specific phobia, social anxiety disorder, panic disorder). These studies have been reviewed by Shin, Rauch, and Pitman (2005a, 2006). In brief, studies have found a reciprocal relationship between medial prefrontal cortex and amygdala function among PTSD patients. For example, in one study employing positron emission tomography and a script-driven imagery paradigm, regional cerebral blood flow (rCBF) increased in the amygdala for PTSD patients relative to controls when trauma-relevant stimuli were presented, and this measure correlated negatively with rCBF changes in the medial frontal gyrus (Shin et al., 2004). Further, PTSD patients have been shown to exhibit exaggerated amygdala responses to fearful versus happy facial expressions relative to controls on measures of blood oxygenation level-dependent signal changes recorded with functional magnetic resonance imaging (Shin et al., 2005b). Here again, diminished medial prefrontal cortex responsivity was observed among PTSD patients.

Wright, Martis, McMullin, Shin, and Rauch (2003) did not find similar amygdala hyperresponsivity to emotional faces in subjects with small animal phobics. Perhaps more noteworthy, symptom-provocation studies with specific phobics have not found amygdala hyperresponsivity to phobic stimuli (Fredrikson et al., 1993; Fredrikson, Wik, Annas, Ericson & Stone-Elander, 1995; Rauch et al., 1995). Wright et al. hypothesized that the absence of responsivity among specific phobics might follow from the innate origins of the associated responses, thereby circumventing any need for amygdala-based learning mechanisms (e.g., fear conditioning).

Other findings are of interest. Increased amygdala activity has been noted among social phobics (Straube, Mentzel & Miltner, 2005) in response to emotional (e.g., angry, happy) and neutral facial expressions, whereas decreased amygdala activity has been found among OCD subjects (Cannistraro et al., 2004). These findings suggest that different anxiety disorders may be associated with different neurocircuitries, although the extent of diagnostic specificity has yet to be established.

Nutt and Malizia (2004) suggested that neuroimaging techniques may serve as a cornerstone in furthering our understanding of the pathophysiology of PTSD. They explained, “Should a marker of brain structure or function definitely be correlated with PTSD, then it is feasible that neuroimaging could be used to identify persons at high risk or to confirm the diagnosis of this disorder ” (p. 16). Several considerations indicate that current research has not yet brought us to the goals envisioned by Nutt and Malizia. First, not all studies demonstrate a pattern of increased amygdala activity among PTSD patients (Phan, Britton, Taylor, Fig & Liberzon, 2006), perhaps reflecting differences
in research methods, patient characteristics, or both. There also remains the possibility that neuroimaging findings represent pre-incident vulnerability factors rather than post-incident consequences of trauma. A study on this question, similar in strategy to research on reduced hippocampal volumes (Gilbertson et al., 2002), is underway (personal communications, L. Shin, July 21, 2006; August 27, 2007). It also is necessary to study additional groups to demonstrate if neuroimaging findings are specific to PTSD. For example, if increased amygdala activity is not associated with specific animal phobias hypothesized to have a partly genetic origin (Wright et al., 2003), might such activity still occur among trauma victims who develop specific phobias after a traumatic event (e.g., learned fear of driving resulting from a motor vehicle accident)? Neuroimaging studies on PTSD and non-PTSD individuals also need to assess individuals in the immediate aftermath of a traumatic event. Thus far, most studies have assessed chronic PTSD patients, leaving room for multiple confounds including medication history and alcohol abuse. Finally, if PTSD symptoms arise largely from an individual’s inability to cope with the meaning of an event (Janoff-Bulman, 1992), associated losses, and cognitive appraisals (Ehlers & Clark, 2000), then the significance of neuroimaging findings may be largely correlational rather than causal (Uttal, 2001).

3.4. Psychophysiological reactivity

Studies have shown that PTSD individuals, as compared with non-PTSD comparison participants, evidence heightened psychophysiological reactivity to trauma relevant cues (Blanchard, Kolb, Gerardi, Ryan & Pallmeyer, 1986; Blanchard, Kolb & Prins, 1991), and to individually tailored script-driven imagery (e.g., Orr, Pitman, Lasko & Herz, 1993; Pitman, Orr, Forgue, de Jong & Claiborn, 1987; Shalev, Orr & Pitman, 1993). In a large study employing more than 1,000 subjects (Keane et al., 1998), a cross-validated formula using physiological measures (e.g., heart rate, skin conductance) correctly identified PTSD and non-PTSD patients at a rate of .64 (assuming a PTSD base rate of .55), indicating somewhat better than chance rates of classification. More recently, a study with monozygotic twins discordant for combat exposure indicated that psychophysiological reactivity represented an acquired response among PTSD diagnosed patients, rather than a genetic vulnerability factor (Orr et al., 2003).

Baldwin, Williams, and Houts (2004) observed how psychophysiological findings have been taken by some authors to demonstrate a “biological signature of PTSD.” For example, Pitman, Orr, Shalev, Metzger, and Mellman (1999) argued that psychophysiological research on PTSD patients had served an important function in dispelling prejudices against the validity of that diagnostic construct. Reflecting the import of this research, the symptom criterion of “physiological reactivity” was added to PTSD’s criterion set in the DSM-III-R (APA, 1987). Yet, the significance of psychophysiological reactivity among many PTSD patients is tempered by the finding that upwards of 40% are not over reactive (Orr, McNally, Rosen & Shalev, 2004). Orr et al. considered a number of hypotheses to explain this limiting finding, although no single explanation was compelling. In a re-analysis of data from several studies, Orr et al. found that physiologic responders and non-responders did not differ significantly in levels of imagery vividness, reported arousal, or negative emotional experiences while recalling traumatic events. These findings bring to mind Lang’s (1968, 1979) early reports on desynchrony among response modalities with phobic subjects.

Even when psychophysiological reactivity occurs, the finding needs to be understood within a broad context of human emotion and arousal. For example, psychophysiological reactivity to anxiety related stimuli has long been observed among individuals with specific phobia (Cook, Melamed, Cuthbert, McNeil & Lang, 1988; Lang, 1968; Weerts & Lang, 1978), with reactivity declining upon successful treatment (Borkovec, 1974). Thus, psychophysiological reactivity to emotionally upsetting stimuli is not specific to PTSD.

3.5. Traumatic memory

Controversies surrounding the concept of traumatic memory speak to what Young (2004b) has termed the “inner logic” of PTSD. Here, Young referred to assumptions regarding underlying mechanisms that account for the manifestation of symptoms after a traumatic event. In this context, several authors have hypothesized that traumatic memories are not merely symptoms that reflect the presence of disorder (e.g., intrusive thoughts, flashbacks). Instead, the formation of traumatic memories represents the principal mechanism or underlying process that drives the clinical syndrome. van der Kolk (e.g., van der Kolk, 1994; van der Kolk & Fisler, 1995) has been one of the major proponents of this inner logic, suggesting that dissociation of traumatic memories is the central pathogenic mechanism that gives rise to the symptoms of PTSD. Throughout his writings, van der Kolk related PTSD symptoms to the “uniqueness of
traumatic memories” (e.g., van der Kolk, 1996, p. 282). Elzinga and Bremer (2002) similarly proposed that PTSD is a “disorder of memory” (p. 12), while Brewin (e.g., Brewin, 2001; Brewin, Dalgleish & Joseph, 1996) proposed a dual representation theory of memory that posits separate systems or “formats” to account for PTSD symptoms.

It is not possible for one section of a review paper to adequately cover all of the hypotheses and debates surrounding traumatic memories. Entire books on the topic have failed to quell dissenting views. Loftus and Ketchum (1994) wrote on the myth of repressed memory, only to be followed by Pope and Brown (1996) who accepted the reality of repressed memories and discussed their proper assessment and treatment. McNally (2003b) provided a comprehensive review and analysis of memory research, and found little, if any, support, for the notion that traumatic memories are unique, remarkably unusual, or immutable. Others have similarly argued that traumatic memory is not special (Shobe & Kihlstrom, 1997). In contrast, Brown, Schefflin and Hammond (1998b) published a comprehensive review of the literature and arrived at very different conclusions.

On occasion, the evidence used in support of unique or distinctive memory processes has been weak or methodologically flawed. For example, in support of traumatic amnesia, Brown et al. (1998a,b) cited a case study (Dollinger, 1985) in which two boys had no memory of an event that led to their friend’s death. McNally (2003b) pointed out that the two surviving boys had been knocked unconscious by the same flashes of lightning that killed their friend, thereby providing a more parsimonious and physical explanation for the boys’ amnesia. Yet not all data are easily dismissed, and some authors continue to accept the reality of psychogenic amnesia (Kopelman, 2002) and the proposition that traumatic memory is special (Nadel & Jacobs, 1998). Furthermore, van der Kolk (1996) has questioned the relevance of many experimental studies that challenge the uniqueness of traumatic memories. He observed that traumatic memories are difficult to study because the profoundly upsetting emotional experiences that give rise to PTSD cannot be approximated in a laboratory setting. By this logic, entire research programs can be dismissed as largely irrelevant to the underlying mechanisms of PTSD (e.g., the Deese–Roediger–McDermott memory illusion paradigm, directed forgetting strategies, emotional Stroop). Similarly dismissed are experimental attempts in the laboratory to assess hypotheses concerning the fragmentation of traumatic memories.

At the heart of debates on the uniqueness of traumatic memory are several hypotheses, sometimes competing and sometimes complementary, but never fully supported (Bernsten, 2001). The various hypotheses cover diverse outcomes ranging from indelible memories that cause morbidity because of their unrelenting presence, to the opposite idea of repressed memories that exert an equally powerful influence without awareness. van der Kolk (1994) suggested that traumatic memories, unlike ordinary recall, could last forever, “timeless and unmodified by further experience” (p. 261). The notion that traumatic memories produce lasting impressions (e.g., flashbulb memories) not subject to decay has failed to receive consistent support (Talarico & Rubin, 2003; Winograd & Neisser, 1992). The hypothesis that trauma leads to splintered or fragmented memories, an idea that on its surface appears opposed to flashbulb memories, also has been challenged in studies that analyze trauma narratives (e.g., Bernsten, Willert & Rubin, 2003; O’Kearney & Perrott, 2006; Rubin, Feldman & Beckham, 2004; Zoellner & Bittenger, 2004). Moreover, at least some of this research is difficult to interpret because of potential confounds. Zoellner and Bittenger observed that measuring purported fragmentation of memory by the coherence of a patient’s narrative is confounded by verbal disruptions associated with emotional arousal. Deducing the occurrence of repressed memories from an individual’s failure to recall confuses different constructs (e.g., normal forgetting vs. repression), and is confounded by an individual’s willingness to report recollections (Femina, Yeager & Lewis, 1990).

Related areas of research, including the hypothesis that PTSD is related to underlying dissociative processes, are beyond the scope of this paper, except to refer the reader to key papers and to note that here again there is equivocal support at best (e.g. Bryant, 2007; Bryant & Harvey, 1997). Thus, dissociation is neither necessary for the development of PTSD, nor is it a specific predictor of the disorder (Barton, Blanchard & Hickling, 1996; Holen, 1993; Marshall & Schell, 2002). Like many other attempts to find a specific mechanism underlying PTSD, dissociation lacks sufficient sensitivity (i.e., most individuals diagnosed with PTSD do not report significant dissociative reactions) to serve as a useful explanatory construct (Bryant, 2007). Similar concerns, and mixed and largely non-supportive findings, apply to the hypothesis that PTSD patients are characterized by a distinctive avoidant encoding style (e.g., McNally, Metzger, Lasko, Clancy & Pitman, 1998; Zoellner, Sacks & Foa, 2003).

The memory wars continue with the occasional call for middle ground (Ost, 2003). Yet after decades of research and heated debate, no hypothesis concerning the distinctiveness of traumatic memories has received sufficient support to validate the PTSD construct as unambiguously separable from mood and other anxiety disorders. Furthermore, and perhaps most importantly, there is no compelling explanation for how distinctive memory processes hypothesized to
result from only extreme, if not horrific, life-threatening events can account for a clinical syndrome that has been shown to result from single episode non-traumatic events (see earlier review of Criterion A).

4. Discussion

The philosopher of science William W. Bartley (1962) maintained that incisive and informed criticism is the ideal route to scientific progress. By rooting out errors in our web of beliefs, constructive criticism allows us to winnow away erroneous assumptions. In a pursuit to accomplish this goal in the case of PTSD, we have found that most every core assumption underlying the diagnostic construct has met with questionable support, if not falsification. This observation does not imply that the sprawling research program surrounding this construct has done little to further scientific knowledge. To the contrary, the construct of PTSD has been of considerable heuristic value in generating research and expanding empirical knowledge. Through this research we have learned more about the characteristics of traumatic and life-altering events, their impact on individuals, and their transaction with pre-incident vulnerability factors, peritraumatic variables, and post-incident buffering supports. It might be argued that research in the field of general stress studies would have produced equivalent results without the PTSD construct, but this position ignores hypotheses regarding processes and outcomes considered special to the more circumscribed field of traumatic stressors. Clearly, the PTSD literature has taught us a great deal.

It is therefore ironic that research whose impetus stemmed primarily from the creation of PTSD has provided a database that challenges most every assumption on which the construct was based. PTSD was originally believed to follow only specific types of traumatic events, but this assumption has been disconfirmed. Traumatic events were generally believed to be the largest contributor to clinical outcome, but this is not the case. Evidence for a dose–response relation between stressors and symptoms has been inconsistent and equivocal. PTSD is defined by symptom criteria that overlap substantially with extant conditions, such as depression and better established anxiety disorders (e.g., specific phobia). Concerted efforts to identify distinct pathogenic processes that underlie PTSD, including research on neuroanatomy, psychophysiological markers and memory processes have met with mixed results, if not outright failure. In short, research has called into question most if not all of the construct-to-manifest indicator nodes within the nomological network surrounding PTSD.

4.1. Caveats

In coming to the present assessment of the PTSD construct, we should emphasize that we have not created various straw-person issues, only to tear them down. Rather, we have critically examined each of the issues raised in the literature by those who have attempted to establish the construct validity of the diagnosis. Further, our analysis does not imply that the construct validity of PTSD rests on any single hypothesis (e.g., evidence of a specific etiology) or on a specific finding (e.g., fragmentation of traumatic memories). To the contrary, there are multiple sources of evidence for the distinctiveness of a diagnosis (e.g., etiologic agent, clinical syndrome characteristics, underlying pathogenesis). Our conclusion that research has failed to support the PTSD construct is based on the accumulated weight of negative evidence, and the failure to convincingly establish any of the key links in the construct’s nomological network.

Of course, it is considerably beyond the scope of any review to evaluate all of the validational issues contained among the 13,774 manuscripts that currently comprise the PTSD literature (search words “posttraumatic stress disorder,” PsycINFO, September 1, 2007). Our review, like any other of a large and sprawling body of literature, is necessarily selective. Further, we are cognizant that others have come to the conclusion that PTSD is a valid diagnostic entity (e.g., IOM, 2006), and that many in the field adopt this position. Nevertheless, because we have taken pains to evaluate large and representative studies of the PTSD construct and because we are unaware of any consistently replicated evidence that countervails our conclusions, we are confident that this review accurately reflects the thrust of the extant scientific literature. In contrast to our efforts, the IOM report entirely disregarded the Criterion A “problem” (Weathers & Keane, 2007), and cited only a single study to conclude that PTSD can be distinguished from other disorders. Moreover, we should bear in mind that negative (potentially falsifying) findings should generally count more than positive (potentially confirming) findings when evaluating scientific theories (Meehl, 1978), including the set of “miniature theories” that comprise a construct’s nomological network. As a consequence, findings in this review that raise serious questions concerning most or even all of the core strands of the PTSD nomological network cannot be casually dismissed as constituting “only a subset of the PTSD literature.”
While attempting to provide a representative and scientifically balanced evaluation of the PTSD literature, we are cognizant that there are large and impressive bodies of literature on PTSD not covered in this review. These areas include, but are not limited to, the use of psychometric data (e.g., MMPI/MMPI-2) in detecting PTSD (e.g., Wilson & Keane, 2004); PTSD’s relationship to physical health (e.g., Schnurr & Green, 2004; Schnurr & Jankowski, 1999) and employment status (Fairbank, Ebert & Zarkin, 1999); additional neuroendocrine and neuroanatomical correlates (e.g., Vasterling & Brewin, 2005; Yehuda, 2006); cognitive processing deficits (e.g., Isaac, Cushway & Jones, 2006); and cross-cultural manifestations of the disorder (e.g., Marsella, Friedman & Spain, 1996).

Still, a compelling argument can be made that these areas of study are marked by methodological limitations and ambiguities similar to – and perhaps even more serious than – those we have reviewed, including confounds due to comorbidity, inadequate levels of diagnostic sensitivity and specificity, and unknown or questionable incremental validity. For example, repeated efforts to identify a distinctive MMPI/MMPI-2 “trauma profile” and/or subscales specific to PTSD (Keane, Malloy & Fairbank, 1984; Lyons & Wheeler-Cox, 1999; Munley, Bains, Bloem & Busby, 1995; Wilson & Walker, 1990) have been largely unsuccessful (Graham, 2006; Greene, 2000; Miller, Goldberg & Streiner, 1995; Wetzel et al., 2003; Wise, 1996). In addition, although PTSD is clearly related to adverse health consequences (Schnurr & Green, 2004), it is not clear that this association is specific to PTSD after taking overlapping mood and anxiety disorders into account. Therefore, nettlesome questions regarding the incremental validity of the PTSD diagnosis are again unresolved. Yet another illustration of this point can be found in a comprehensive review on episodic memory deficits, in which Isaac et al. (2006) concluded there was little evidence to suggest that the cognitive profile of PTSD differed from that of other anxiety or depression disorders.

Our conclusion that findings have generally failed to support core assumptions underlying PTSD should not be taken to imply that construct problems are unique to this diagnosis. Much has been written on systemic problems with psychiatric nomenclature since the advent of the DSM-III (e.g., Beutler & Malik, 2002; Houts, 2000; McHugh, 1999), and on how mental disorder is best defined (e.g., Houts & Follette, 1998; Kirmayer & Young, 1999; Lilienfeld & Marino, 1995; Spitzer, 1999; Wakefield, 1992, 1999a,b). Most, if not all, diagnoses contained in various editions of the DSM lack a specific etiology, a pathognomonic symptom or constellation of pathognomonic symptoms, and a distinct pathogenesis. At the same time, it can be observed that foundational nosological issues and unresolved controversies surrounding the validity of diagnoses such as major depression, bipolar disorder, schizophrenia, specific phobia and obsessive–compulsive disorder have not grown over various editions of the DSM, nor have the operational criteria for these disorders necessitated substantial revisions (Spitzer et al., 2007). Further, there is no evidence to suggest that a diagnostic construct like depression or schizophrenia is better, or just as well accounted for by another category within the DSM.

There is also one crucial respect in which PTSD established for itself a higher standard than extant diagnoses. PTSD, in sharp contrast to virtually all other psychiatric diagnoses (e.g., schizophrenia, major depression, panic disorder, GAD, OCD), has been modeled explicitly around the core assumption of specific etiology (Meehl, 1977). As a consequence, it must meet an additional burden not demanded of other diagnoses, namely, the demonstration of a strong and specific causal linkage between stressors and symptoms. As we have noted in this article, the evidence bearing on this linkage is equivocal at best. Surprisingly, the IOM report (2006) bypassed this critical requirement in their analysis of PTSD’s construct validity.

Most important among our caveats is the need to emphasize that an acknowledgement of quandaries besetting PTSD in no way denies the intense emotional pain and suffering experienced by many trauma victims. This is a particularly important issue at a time when American troops are returning from combat in Iraq and valid mental health needs must be addressed. Many years ago, Breslau and Davis (1987) were similarly careful to address this point when they challenged the validity of Criterion A and the assumption that a distinctive set of events was associated with a distinct clinical syndrome. They stated:

As we attempt a critical understanding of the DSM-III definition of PTSD, we do not deny the clinical reality of the psychological suffering and psychiatric symptoms that are the sequelae of traumatic experiences. Our concern is with the validity of the set of explicit criteria that define PTSD in DSM-III. (p. 255)

Despite this attention to clinical concerns, Lindy, Green, and Grace (1987) wondered if Breslau and Davis had suggested that perhaps “the systematic study of survivors of catastrophe [should] once more be ignored by our field.” (p. 271). Lindy et al. called such a suggestion “actively harmful on ethical grounds,” apparently neglecting the point that Breslau and Davis had never advanced this position.
The sentiment that those who objectively assess the validity of the PTSD construct are “arguing” (Kilpatrick, 2006), or somehow challenging or abandoning victims of trauma (e.g., Lindy et al., 1987) is misguided. The question of concern when assessing the validity of PTSD is not the political, economic, or social protection of victims, all of which are profoundly important but logically separate matters from PTSD’s construct validity. Rather, the question of concern is the appropriateness of applying a diagnostic construct to individuals and their emotional states in the aftermath of trauma. In a similar vein, investigating the validity of alternative conceptualizations and pursuing accurate information about trauma is ultimately the best way to advocate for victims (Rosen & Frueh, 2007).

4.2. Proposals for the DSM-V

In 2004, the American Psychiatric Association launched the DSM-V Prelude Project (APA; http://www.dsm5.org). Initial steering committees began to meet the next year in preparation for the development of the DSM-V. The current but tentative timeline provided by APA’s Prelude Project is for work groups to be appointed shortly, with the publication of the DSM-V in or around 2011. Among the hundreds of diagnoses now in the DSM, PTSD is one certain to receive intense debate. The PTSD work group will have numerous challenging issues to tackle, many of which were considered by the framers of DSM-IV (Davidson et al., 1996; Rosen et al., 2008; Spitzer et al., 2007). What definition of the stressor criterion (Criterion A) can best direct research toward the question of a specific etiology? Is another set of changes in the definitions and groupings of symptom criteria required? Should PTSD be classified as an anxiety disorder or a dissociative disorder, or should there be an entirely new category of stress- or trauma-related psychiatric illnesses within which PTSD would be included?

Of course, one option for the future of PTSD in the DSM-V is to leave its criteria set alone. Although this approach might have the benefit of providing continuity, it would fail to address all the troubling questions concerning construct validity raised by years of research. Therefore, many commentators who have discussed PTSD’s future have proposed changes in its defining criteria. Change, however, is a tricky matter, because alternative proposals tug in competing directions. For example, Resick (2004) reviewed the many problems with PTSD as currently operationalized, while focusing primarily on the symptom criteria. She reviewed research suggesting that (a) diverse emotional reactions occur after trauma, including grief, shame, disgust, and anger and (b) non-Western cultures give different expression to traumatic loss. Resick proposed an expanded definition of PTSD that moved the construct out of the supra-heading of “Anxiety Disorders” into a new classification of “Stress-Related Disorders.” In this new supra-heading would be listed the adjustment disorders, possibly Acute Stress Disorder, PTSD, Traumatic Grief, and Dissociative Disorders. Further, according to Resick’s proposal, the symptom criteria would be broadened to include a greater range of adverse emotions. A possible advantage of this approach is that it would help restore PTSD to its origins in the broader field of stress studies (e.g., Cannon, 1929, 1935; Dohrenwend & Dohrenwend, 1974a,b; Selye, 1936, 1946).

Resick’s analysis is accurate in pointing out that a host of emotional reactions follow adverse events, including anger (Novaco & Chemtob, 2002), guilt (Kubany, 1994; Kubany et al., 1996), shame (Andrews, Brewin, Rose & Kirk, 2000), and disgust (Foy, Sipprelle, Rueger & Carroll, 1984). Each of these emotional states could be joined with the PTSD construct or split off to create its own disorder. For example, an extensive literature exists on the relationship of disgust reactions to anxiety processes (e.g., McKay, 2002; Sawchuk, Lohr, Westendorf, Meunier & Tolin 2002). Further, studies on PTSD have demonstrated a link between grotesque imagery (an event characteristic associated with disgust reactions) and severity of posttraumatic reactions (Green, 1990; Green, Grace, Lindy, Gleser & Leonard, 1990; McCarroll, Ursano, Fullerton & Lundy, 1996). If the study of disgust should shift from the paradigm of behavioral analysis (identifying event-response relationships) to the taxonomic framework of the DSM (specifying categorical disorders) there would be two options. One option, as suggested by Resick, would be to expand the criteria for PTSD and accommodate such emotional dimensions as disgust. This would represent a new form of criterion creep, and raise the concern that PTSD might become a cultural narrative that medicalizes a broad range of reactions to adversity (Rosen, 2004; Shephard, 2004; Summerfield, 2001, 2004). Of course, if an expanded criteria set were rejected, then the other option for dealing with additional negative affective states would be to create Posttraumatic Disgust Disorder.

The idea of creating new disorders based on the PTSD syndromic model has been advanced by several individuals. Linden (2003) proposed the diagnosis of Posttraumatic Embitterment Disorder (PTED) to further the study and treatment of individuals who suffer severe embitterment after adverse events. Vandervoort and Rokach (2004) discussed Posttraumatic Relationship Syndrome (PTRS). Prigerson and colleagues suggested that the DSM-V should include a diagnosis to cover Traumatic Grief Disorder (Prigerson & Jacobs, 2001) or Complicated Grief (Lichtenthal,
Cruess & Prigerson, 2004). Prolonged Duress Stress Disorder has been proposed for individuals who experience posttraumatic reactions without any specific stressful event (Waddington et al., 2003). This expanding list of disorders covers an increasing array of human reactions to adversity without resolving any of the underlying problems that have befallen the original construct of PTSD.

On the other side of the debate for changes in PTSD’s defining criteria are proposals for a tighter (more restrictive) definition. As previously discussed, Foa et al. (1995) proposed that the diagnosis of PTSD might be limited to individuals who evidence emotional numbing, with other symptomatic individuals diagnosed with an alternative anxiety disorder. Within this framework, PTSD would maintain its historical roots within the Anxiety Disorders, four symptom clusters would be identified among extant symptoms, and increased emphasis would be placed on the emotional numbing criteria. This proposal might enhance the discriminant validity of PTSD from other anxiety disorders.

A recent proposal by Spitzer et al. (2007) recommends tightening the definition of Criterion A so that only events directly experienced would qualify individuals to be considered for the PTSD diagnosis. Tightening of the symptom criteria was also suggested by Spitzer et al., with the recommendation to drop symptoms that overlap with other diagnoses (e.g., removing difficulty concentrating and difficulty sleeping, which are redundant with criteria for depression). In this manner, Spitzer et al. hope to contain problems associated with criterion creep and the non-specificity of the PTSD syndrome.

There are inherent risks with proposals that change how PTSD is operationalized. A change in defining criteria opens the doors for new ways that findings and underlying assumptions may not fit. New assessment instruments would be required and studies would need to show how extant findings fit with the redefined construct. Proposed changes in the PTSD construct also fail to solve many of the fundamental issues discussed in this paper. For example, a tightening of the symptom criteria with greater emphasis on the role of emotional numbing does nothing to resolve fundamental problems with Criterion A. Resick’s proposal to move PTSD to a supra-heading for stress-induced disorders extends the assumption of specific etiology to its logical limits, with all the problems that attend this assumption. New models for defining PTSD also do not solve the problem of criterion creep, and it remains unclear what creative permutations might arise within these models.

4.3. Should PTSD be moved to Appendix B?

However heretical it may sound, it is logical to ask if PTSD has performed so poorly that it should be dropped from the manual. Such action is not without precedent in the area of stress studies. For example, DSM-I (APA, 1952) included “Gross Stress Reaction” (p. 40), a category that disappeared with the publication of the DSM-II (APA, 1968). More recently, it has been proposed that Acute Stress Disorder, a construct first introduced in DSM-IV be dropped from future editions of the DSM (Bryant, 2007; Spitzer et al., 2007). After 25 years of research and over 13,000 published studies, should PTSD receive a similar fate?

When considering the possibility of removing PTSD from DSM-V, there are several social and historical considerations. First, it seems unlikely that any effort to eliminate PTSD would succeed even if data unambiguously supported the move. As observed by the English historian Ben Shephard,

If “trauma” could now be broken up into its constituent parts, it would return to its social contexts and be de-medicalized. But for that to happen, psychiatry would have to surrender ground, and history teaches us that such acts of professional self-denial are indeed rare. Besides, it is now too late. Trauma has been vectored into the wider society by the law and the media. (pp. 57–58)

Shephard’s observation that PTSD is now entrenched in the psychiatric nomenclature must be qualified by an historical lesson: posttraumatic syndromes can be overgeneralized and abused, eventually leading to their downfall. War-related syndromes provide the best known examples of this pattern (Shephard, 2001), but other instances can be found. Killen (2003) offered an historical analysis of German telephone operators in the early 1900’s that revealed (a) creation of a compensation system for those who succumbed to the pressures of their work, (b) growth of compensation claims overburdening the system, and (c) development of the view that the claims process itself, rather than work, was the true pathogen. In the end, laws were changed, the problem was redefined, and the medical syndrome that protected victims was eliminated. Might a similar fate eventually befall PTSD if excesses and criterion creep continue unrestrained, or if unfounded claims of disability (Frueh et al., 2005) and courtroom cases (Rosen, 1995)
cause a backlash against the construct? It would be ironic if PTSD met its demise at the hands of social and political forces, as some have suggested that these were the sources of influence from which the construct emerged in the wake of the Vietnam War and its aftermath (Scott, 1990; Shephard, 2001; Young, 1995).

Returning to the question of whether PTSD should be removed entirely from the DSM, we think not. Despite its many failings, the construct is linked to numerous testable hypotheses and clinical concerns that warrant further investigation. This being the case, we examine one remaining option for the future of PTSD in the DSM. An often overlooked section in the DSM-IV is Appendix B. This appendix provides criteria sets and axes for further study. In other words, experimental criteria sets not yet established as valid clinical syndromes. In hindsight, we believe PTSD should have been listed in such a section when introduced in 1980. However, it was not until DSM-III-R (APA, 1987) that an appendix was introduced to provide for, “Proposed Diagnostic Categories Needing Further Study.” If PTSD had been moved to this appendix in DSM-III-R, then research could have been encouraged without the same susceptibility to expansion and premature reification of the construct. We believe this is still a viable option.

If PTSD were redefined as an experimental criteria set it would be possible to list competing proposals in an effort to encourage systematic research. Alongside the current criteria set that has been in use since DSM-IV (option “a”), there could be alternative proposals that might include: (b) an expanded criteria set with additional negative affect states and broader definitions of Criterion A; (c) a restricted criteria set that emphasized emotional numbing; and/or (d) the criteria set recently proposed by Spitzer et al., (2007). If over time, one criteria set performed best in distinguishing PTSD as a valid construct separable from extant conditions, then that description of the clinical syndrome could be returned to the DSM as a recognized Axis I disorder.

4.4. Concluding thoughts

Regardless of how PTSD appears in the DSM-V, we offer several observations and recommendations. First, various research programs have too often pursued a pet hypothesis without linkage to other relevant variables. For example, several investigators have attempted to tie neurobiological models that are hypothesized to underlie PTSD with the particular features of traumatic memory (e.g., Brewin, 2001; Charney, Deutch, Krystal, Southwick & Davis, 1993; Elzinga & Bremner, 2002; van der Kolk, 1994). The dots are connected by observing findings in diverse areas (e.g., cortisol, hippocampal volumes, trauma narratives) without conjointly assessing these multiple processes within the same study. The field has reached a level of maturity wherein broader collaborations may be productive. For example, we do not know whether individuals with strong dissociative tendencies possess fragmented memories for high magnitude traumatic events or whether these same individuals have low cortisol levels or small hippocampal volumes. Nor do we know whether increased amygdala activity occurs most prominently among PTSD patients with high levels of emotional numbing. If such diverse findings converged to identify a specific patient subgroup in the large universe of individuals currently meeting PTSD criteria, then perhaps a disorder in nature could be identified and the distinctiveness of a trauma syndrome affirmed.

Second, we encourage research on that interesting subgroup of individuals who manifest severe posttraumatic reactions and meet PTSD symptom criteria, but who have not experienced a traumatic event. Such research would examine the characteristics of individuals who all ostensibly share a strong (genetic) propensity to developing PTSD symptoms (although in this case such symptoms would not be truly “post traumatic”) even in the absence of strong stress. It could turn out that an identified diathesis is nothing more than the “usual suspect” of high trait neuroticism or negative emotionality (NE), which correlate highly with PTSD (e.g., Davidson, Kudler & Smith, 1987; Kuhne, Orr & Barage, 1993), but that finding in itself would be informative. Alternatively, it remains possible that the investigation of this poorly understood subgroup would identify a specific diathesis to PTSD symptoms above and beyond NE and other anxiety disorders.

In closing we take note that, despite considerable issues and controversies, PTSD in some reconstituted form may yet achieve scientific status as a diagnosis that holds compelling evidence for construct validity. While we await such evidence, an important task for those shaping PTSD in the DSM-V will be to maintain a reasonable perspective through the provision of cautionary statements and guidelines. There is precedent for this approach. Text in the DSM-IV introduced a cautionary guideline for clinicians and researchers to rule out malingering in cases involving financial remuneration and forensic determinations (APA, 1994). Although not always heeded, this guideline directs the attention of professionals to an important concern while safeguarding the accuracy of the PTSD database (Rosen, 2006).
Additional qualifications to the DSM’s handling of PTSD text may be similarly helpful. Based on current findings, the DSM-V should emphasize that typically, if not always, only a minority of individuals develop reactions of sufficient breadth, severity, and duration to warrant the diagnosis of PTSD. Text in the DSM-IV is not explicit on this point. Qualifying statements should emphasize that Criterion A stressors are not sufficient to bring about the full symptom picture of PTSD, reminding the clinician that a patient’s presenting problems requires a detailed analysis of causality rather than simply listing a single event (McHugh & Treisman, 2007). Additional text could caution clinicians who must distinguish relatively normative reactions to stress from the symptoms of a dysfunction within the organism (Spitzer et al., 2007).

Finally, rather than conceptualizing PTSD as a taxon (i.e., category in nature; Meehl & Golden, 1982), it may be more fruitful and scientifically supportable to consider PTSD as encompassing a broad range of possible reactions to adverse events. These reactions are in turn influenced by multiple dimensionally distributed factors: an individual’s pre-incident risk characteristics and life circumstances; event characteristics and the individual’s appraisals of those characteristics; an individual’s reactions during the event and appraisals of those reactions; direct consequences of the event in terms of loss or injury; impact of the event on the individual’s adjustment and core cognitive schemata; and post-event factors that ameliorate or reinforce the maintenance of initial reactions, the development of new ones, or both. If we have learned anything from the long history of posttraumatic studies, and the shorter history of PTSD, it is that multiple factors and their complex interrelations result in variable outcomes that are unlikely to be explained by a single disorder in nature.

References


