Events, Reactions, and Behaviors: Assessment of Posttraumatic Stress Disorder

By Erienne R. Weine

B.S. in Psychology, University of Arizona
B.A. in Philosophy, University of Arizona
M.A. in Psychology, Northeastern University

A dissertation submitted to

The Faculty of
the College of Science of
Northeastern University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

August 5th, 2015

Dissertation directed by

Nancy S. Kim
Associate Professor of Psychology
Dedication

I dedicate this dissertation to my family.
Acknowledgements

I am deeply grateful to the following people for making this journey possible. Most of all, I wish to thank my advisor Nancy Kim. I honestly could not ask for a better mentor. Thank you for years of your careful guidance, time, attention, feedback, kindness, and encouragement, from which I have grown and benefitted immensely. You have taught me in innumerable ways and I will forever be grateful.

Thank you to the Cognition Area for playing a central role in my education. Our many thought provoking discussions made me a stronger thinker and researcher. Your generous and honest feedback allowed me to sharpen my skills as a speaker. Especially I want to thank John Coley for mentoring me throughout my time as a graduate student. Thank you Joanne Miller for the many opportunities to learn about higher education administration, for your guidance in that endeavor, and for the opportunity to teach. Thank you my dear Cognition Area graduate student friends for years of shared understanding, support, encouragement, and conversation: Nicole Betz, Susan Dalrymple, Amy DiBattista, Amanda Dupuis, Cole Eidson, Maureen Gillespie, Mason Jenkins, Darrell Penta, Nadya Vasilyeva, Yian Xu, Jennelle Yopchick, and Xu Zhao.

Thank you to the wonderful Causal Cognition Lab. Thank you Nancy Kim for creating such a positive and productive lab. To my research sisters and brother: Jennelle Yopchick, Susan Dalrymple, and Mason Jenkins, thank you for being my role models, sounding board, and support team. To the many Research Assistants with whom I had the pleasure of working in the Causal Cognition Lab: this research could not have been accomplished without your dedication, engagement, and willingness to ask questions. Thank you!
I also wish to thank the many other wonderful people in the Psychology Department who mentored me: Allison Baker, Iris Berent, Marty Block, Heather Brenhouse, Judith Hall, Neal Pearlmutter, Adam Reeves, and Nancy Snyder. To all of the graduate students with whom I’ve interacted during my time here, thank you for openly sharing the experience of being a graduate student with me.

Finally, I wish to thank my dear friends and family. Thank you Dave Richters for teaching me about life during and after graduate school, and for your delightful company. Thank you Sam Simmers for sharing a beautiful love with me. You have supported me emotionally, mentally, and in many other ways, especially throughout the past year, and I am deeply grateful. For your unconditional love, thank you to my mother, brother, sister, grandmother, aunts, uncles, and cousins. Finally, I wish to thank my mother for raising me to be a happy, strong, driven, positive, loving, and grateful individual. I am proud to be your daughter.
Abstract of Dissertation

There is an ongoing debate in the mental health field over whether mental disorder diagnosis should take contextual factors into account (Horwitz & Wakefield, 2007). However, it is unclear whether people are actually influenced by context when identifying disorders. Research suggests that clinicians, clinical trainees, and lay people explain away the abnormality of mental disorders in causal context (Ahn, Novick, & Kim, 2003), but whether causal context also leads to fewer diagnoses has not yet been determined. In the Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013), some causal context is required for the diagnosis of Posttraumatic Stress Disorder (PTSD); specifically, the experience of a traumatic event (e.g., killing a civilian in combat). On the other hand, recent survey and interview data indicate that the onset of PTSD symptoms (e.g., exaggerated startle response) is not predicted most reliably by a traumatic event, but rather by a severe internal reaction (e.g., extreme fear; Boals & Schuettler, 2009). Therefore, in the current work, we investigated the extent to which clinicians, clinical trainees, and lay people are influenced by the context of a traumatic event, a severe reaction to an event, or both when diagnosing PTSD.

Our two central questions were (1) whether people identify PTSD more or less often in causal context, and (2) whether they are influenced by the type of context specified by the current version of the DSM (i.e., traumatic event context only; DSM-5, APA, 2013), recent survey research (i.e., the context of an intense reaction to any event; e.g., Boals & Schuettler, 2009), or the previous version of the DSM (i.e., traumatic event context and intense reaction context; DSM-IV-TR, APA, 2000). We also examined several supplementary questions to place the investigation of PTSD diagnosis in broader, appropriate, context. These included (3) whether
these findings would extend to people’s identification of other disordered behaviors (specifically, MDD) and non-disordered behaviors, and (4) whether we could replicate past findings that disordered behaviors seem less psychologically abnormal in causal context across disorders. We also examined (5) whether there were any differences in the judgments of clinicians and clinical trainees on these tasks. Finally, we examined whether clinicians and clinical trainees were (6) more likely to form distorted memories of reactions than events or behaviors, and (7) whether they were specifically susceptible to forming distorted memories of reactions whose strength of negative valence was proportionate to that of the other two elements.

Experiment 1 addressed Questions 1 through 4 in a sample of lay people, and Experiments 2 and 3 addressed Questions 1 through 7 in samples of clinical trainees and licensed clinicians. In all three experiments, participants read about hypothetical people presented in vignettes describing PTSD behaviors, MDD behaviors, and distressed behaviors, within-subjects. The behaviors were presented in one combination of event and reaction context as determined by a 2 (event type: traumatic event, everyday event) x 2 (reaction type: intense reaction, mild reaction) between-subjects design. Participants were asked to provide diagnoses for the vignettes, using an open-ended measure, and were also asked to judge the psychological abnormality of the cases.

Across experiments, we found that the presence of traumatic event context increased the incidence of PTSD diagnoses. Clinicians, clinical trainees, and lay people all identified PTSD more often when PTSD behaviors were presented in traumatic event context than in everyday event context. Reaction context did not influence diagnostic judgments; thus, participants diagnosed PTSD in line with the DSM-5 (APA, 2013). Yet, these context effects held true only for PTSD. MDD behaviors were correctly identified as depression regardless of context;
distressed behaviors were most often identified as describing no disorder at all, also in all combinations of contexts (although less often in the context of a traumatic event in Experiment 2). We replicated past research demonstrating that disordered behaviors seem less psychologically abnormal in causal context, suggesting that perceptions of abnormality and diagnostic judgments do not appear to be similarly influenced by causal context. Moreover, we found no differences in the influence of context on the diagnostic judgments of clinicians versus clinical trainees.

Because we found no evidence that reaction context influenced any of the judgments measured, in Experiment 3, we investigated clinicians’ and clinical trainees’ memories for events, reactions, and behaviors using a standard recognition task. Both groups demonstrated poorer memory for reactions than for events or behaviors. Participants were particularly susceptible to falsely recognizing reactions that were distorted in strength of negative valence to better align with the events and behaviors.

Taken together, our findings suggest that PTSD is selectively identified most often in traumatic event context only, by lay people, clinical trainees, and clinicians. This indicates that all three groups identify PTSD in line with DSM-5 (APA, 2013) specifications, despite the fact that many of them reported having little to no familiarity with this latest version of the manual at the time of the study. Implications for clinical practice, the clinical utility of the DSM system, and the debate over the use of context in diagnosis are discussed.
Table of Contents

Dedication iii
Acknowledgements iv
Abstract of Dissertation vi
Table of Contents ix
List of Tables x
List of Figures xii
Introduction to Dissertation 1
Paper 1: College Students’ Identification of Mental Disorders in Context 5
Interim Discussion 59
Paper 2: Events, Reactions, and Behaviors: Clinical Assessment of Posttraumatic Stress Disorder 61
Appendix A: Full Vignettes 165
Appendix B: College Student Open-Ended Disorder Identification Responses 188
Appendix C: Pilot Study: College Students’ Identification of Mental Disorders Using a Closed-Ended Measure of Diagnosis Judgments 195
Appendix D: Clinical Open-Ended Disorder Identification Responses 203
Appendix E: Supplementary Experiment: Clinicians’ Diagnoses and Judgments of Treatment Need, Distress, Functional Impairment, and Psychological Abnormality 209
List of Tables

Table A1: Partial Sample Vignette – “Recent Past” 47
Table A2: Dependent Variables 49
Table A3: Open-Ended Disorder Identification Responses by Behavior Type 50
Table A4: Number (Percent of Total) of Open-Ended Disorder Identification Response Categories by Behavior Type 52

Table B1: Experiment 1 Partial Sample Vignette – “Recent Past” 144
Table B2: Experiment 1 Dependent Variables 146
Table B3: Experiment 1 Open-Ended Disorder Identification Responses by Behavior Type 147
Table B4: Experiment 1 Mean Ratings for Psychological Abnormality Judgments and Difficulty Understanding Judgments by Context and Behavior Type, with Statistical Significance of Event Type Main Effects from 2x2 ANOVAs 149
Table B5: Experiment 2 Event Recognition Items 150
Table B6: Experiment 2 Reaction Recognition Items 153
Table B7: Experiment 2 Behaviors Recognition Items 155
Table C1: All Open-Ended Diagnosis Responses by Clinicians and Clinical Trainees in Experiment 1 and Experiment 207
Table D1: Appendix E Dependent Variable Text and Instructions 234
Table D2: Non-DSM Event Vignette Type and Text 236
Table D3: PTSD Behaviors Correlation Matrix for PTSD Diagnosis Likelihood 238
Judgments, PTSD Treatment Need Judgments, Distress Judgments, Functional Impairment Judgments, and Psychological Abnormality Judgments

Table D4: Multiple Regression Results for PTSD Behaviors 239

Table D5: MDD Behaviors Correlation Matrix for MDD Diagnosis Likelihood 240

Judgments, MDD Treatment Need Judgments, Distress Judgments, Functional Impairment Judgments, and Psychological Abnormality Judgments

Table D6: Multiple Regression Results for MDD Behaviors 241
List of Figures

| Figure A1: Proportion of PTSD Identifications | 53 |
| Figure A2: Proportion of MDD Identifications | 54 |
| Figure A3: Proportion of “None” Identifications | 55 |
| Figure A4: Likelihood of Disorder Identifications by Event and Reaction Type | 56 |
| Figure A5: College Student Psychological Abnormality Judgments | 57 |
| Figure A6: College Student Difficulty Understanding Judgments | 58 |
| Figure B1: Experiment 1 Clinicians’ and Clinical Trainees’ Reported DSM-5 Experience | 158 |
| Figure B2: Experiment 1 Number of Disorder Identifications by Event Type | 159 |
| Figure B3: Experiment 2 Open-Ended Diagnosis Encoding Task Results | 160 |
| Figure B4: Experiment 2 Manipulation Check Severity Ratings | 161 |
| Figure B5: Experiment 2 Proportion of False Alarms for Events, Reactions, and Behaviors | 162 |
| Figure B6: Experiment 2 Proportion of Hits for Events, Reactions, and Behaviors | 163 |
| Figure B7: Experiment 2 Proportion of False Alarms for False Strength Reaction Items in Proportionate and Disproportionate Reaction and Behaviors Conditions | 164 |
| Figure D1: PTSD Diagnosis Likelihood Judgments for PTSD Behaviors by Event Type and Expertise | 242 |
| Figure D2: PTSD Treatment Need Judgments for PTSD Behaviors by Event Type and Expertise | 243 |
Introduction to Dissertation

Mental disorder is a broad term describing abnormal functioning in an individual such as impairment, statistically rare behaviors, and maladaptive responses to circumstances (Lilienfeld & Marino, 1995). The term mental disorder differs across cultures (Bennett, 2003) and escapes precise definition (DSM, 3rd ed.; DSM-III; APA, 1980; Lilienfeld & Marino, 1995). Nonetheless, a mental disorder can be generally characterized as an abnormal, negative state in an individual. This negative state is caused by distorted psychological processes, may place the individual in danger, and merits alleviation through care or treatment (Bennett, 2011; Kring, Johnson, Davison, & Neale, 2010). In all editions of the Diagnostic and Statistical Manual of Mental Disorders (DSM) since 1980 (DSM-III; APA, 1980), the manual has defined a mental disorder as an abnormal, dysfunctional state in an individual that causes personal distress or disability, may pose further risks of pain, death, loss of freedom, distress or disability, and is not merely a violation of cultural or societal norms.

The DSM itself is a handbook created and published by the American Psychiatric Association (APA) to define mental disorders and categories for use in diagnosis and research in the U.S. and other countries (DSM-5; APA, 2013). The DSM broadly defines clinical diagnosis as assessment of behavioral symptoms that are perceived to be abnormal (e.g. someone experiencing distress and impairment in daily life). In the DSM, mental disorder categories are named, described, and presented alongside prevalence rates and other statistical information. Importantly, the DSM includes instructions for diagnosing each disorder (Blashfield, Keeley, Flanagan, & Miles, 2014). That is, most DSM disorders are each defined by a set of rules: A list of diagnostic criteria and exclusions that specify whether a diagnosis should be made in any given case (DSM-5; APA, 2013).
Clinicians make diagnostic judgments for a client, in part, to identify appropriate treatments that have already been studied for efficacy in groups of similarly suffering individuals (e.g. Ahn, Proctor, & Flanagan, 2009). In the U.S. (Kutchins & Kirk, 1997), clinicians must provide a *DSM* diagnosis for their clients to receive insurance reimbursement for mental disorder treatment. This effectively constrains formal clinical diagnosis to this established framework, which is designed to bring consensus amongst clinicians (Kirk & Kutchins, 1992). However, clinicians do not represent mental disorders as a list of symptoms, the way they are described in the *DSM*, but rather as a theory-based concept in which the symptoms cause and effect one another and differ in importance (Kim & Ahn, 2002). Because of this, clinicians are more likely to diagnose a hypothetical person with symptoms that are causally central to the clinician’s concept of the disorder than a hypothetical person with causally peripheral symptoms (Kim & Ahn, 2002). That is, clinicians do not appear to diagnose mental disorders by identifying any set of diagnostic criteria, as the *DSM* instructs them to do.

Lay assessments of mental disorders are just as important as professional clinical assessments, because lay people are on the front lines of diagnosis (Pescosolido, 2006). Lay people start to formulate ideas about what might be wrong with a family member or friend well before these peers ever see a clinician. Thus, they are often the first to identify a problem and recommend help (Pescosolido, Brooks-Gardner, & Lubell, 1998; Pescosolido & Perry, 2014). Lay assessments of mental disorders are likely indirectly influenced by *DSM*, at least in part, as lay beliefs about mental disorders appear to align closely with expert concepts (e.g., depression), even though it is unlikely that lay people engage with the *DSM* directly (Shaw, 2002). Whooley (2010) similarly found in a qualitative study that clinicians don’t necessarily always follow the
DSM, but rather are prone to use their own clinical judgment. Thus, it is important to better understand how and when both clinicians and lay people make diagnostic judgments.

Posttraumatic Stress Disorder (PTSD) is characterized by a stressful experience that leads to symptoms months later such as flashbacks, a heightened startle reflex, negative thoughts and feelings, impairment in daily functioning, and avoidance of stressor reminders (Diagnostic and Statistical Manual of Mental Disorders [DSM], 5th ed.; DSM-5; American Psychiatric Association [APA], 2013). The DSM defines PTSD in all editions from 1980 (DSM-III; APA, 1980) to the present time (DSM-5, APA, 2013) as a list of diagnostic criteria, including a traumatic stressor, re-experiencing the stressor, numbing, avoidance, and heightened arousal. An estimated 7.8% to 8.3% of adults in the U.S. develop PTSD in their lifetime (Kilpatrick et al., 2013; Gabbay, Oatis, Silver, & Hirsch, 2004; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995). Further, approximately half of people in the U.S. will develop one or more mental disorders in their lifetime (Kessler et al., 2005).

This dissertation examines how people (lay people, clinicians, clinical trainees) make diagnostic judgments about potential cases of PTSD. It is comprised of two papers reporting these experiments, with the intention that this work will ultimately be submitted for publication as two papers. The first paper, intended to target a health communication or health literacy journal, presents a study that examines the mental health literacy of college students. Specifically, the first paper investigated whether college students can identify mental health disorders in context, including PTSD. The second paper intended to target a clinical journal, presents a set of two studies that investigated the context in which clinicians and clinical trainees diagnose PTSD. The papers are supplemented by five appendices that include full vignettes,
more complete descriptions of open-ended diagnoses not included in the manuscript, and an additional two studies I conducted that informed this research.
College Students’ Identification of Mental Disorders in Context

Erienne R. Weine, M.A. & Nancy S. Kim, Ph.D.

1 Department of Psychology, Northeastern University, 125 Nightingale Hall, 360 Huntington Avenue, Boston, MA 02115, U.S.A.

* Corresponding author: Erienne R. Weine. Department of Psychology, Northeastern University, 125 Nightingale Hall, 360 Huntington Avenue, Boston, MA 02115, U.S.A., Tel: 520-603-3624; Fax: 617.373.8714; Email: e.weine@neu.edu.

We thank Tara Armor, Marielle Evangelista Filler, Rachel Foulkes, Dunia Kassay, Britney Leibert, Jordan Nissen, Katharine Quain, Alicia Rogers, and Cara Vincelette for help with checking materials, programming, data collection, and coding.
Abstract

Nearly half (46%) of college students in the U.S. have a mental disorder, and only a fraction of these seek treatment (19%; Blanco et al., 2008). Poor mental health literacy is a major barrier to treatment, and one aspect of mental health literacy is whether people, including older adolescents, can recognize mental disorders (Jorm, 2000; Gulliver et al., 2010). One factor that affects disorder assessment is causal context (Ahn et al., 2003); therefore, we investigated whether college students are able to identify disorders in context. Some mental disorders occur in very specific causal context (e.g., PTSD is caused by a traumatic life-event; DSM-5, APA, 2013), whereas some do not (e.g., major depressive disorder [MDD] need not have a specific cause; DSM-5, APA, 2013). College students aged 18 to 23 learned about cases of hypothetical people exhibiting behavioral symptoms of PTSD and MDD, each presented in the context of a causal life-event. Using an open-ended measure, we asked students to identify disorders for these cases. The presence of a traumatic life-event was sufficient to motivate college students to identify PTSD for either type of disordered behavior, although more so for PTSD behaviors. MDD behaviors, appropriately, were the only ones identified as MDD, and were identified without context. College students appear to be reasonably accurate at identifying these common disorders in context, in hypothetical others. We discuss alternative factors that may help account for college students’ frequent failure to seek treatment.

(239 words)
College Students’ Identification of Mental Disorders in Context

Over half of college students in the U.S. likely qualify for a mental health disorder diagnosis. That is, the prevalence of mental disorders in this population over a 12-month period is about 46%, according to the National Epidemiological Survey on Alcohol and Related Conditions (NESARC; Blanco et al., 2008). The onset of mental health disorders is most typically in late adolescence and young adulthood (Kessler et al., 2007), in part due to major neurobiological, affective, and hormonal changes that span years into older adolescence (Paus, Keshavan, & Geidd, 2008). In addition to navigating the biological challenges of older adolescence, college students are particularly susceptible to developing a mental disorder because they must learn to manage critical academic, social, and emotional skills on their own when transitioning from high school to college (Norton & Brett, 2011). Nevertheless, the vast majority of college students who might have a mental disorder do not seek treatment for their symptoms (81%, Blanco et al., 2008). Further, those who do seek treatment often delay doing so for years (Thompson, Isaakidis, & Hunt, 2008).

Mental Health Literacy

In addition to fear and uncertainty about what to do (Brandli, 1999), poor health literacy is a major barrier to treatment (Jorm, 2000; Kreps, Bonaguro, & Query, 1998; Saunders, 1993). Health literacy is knowledge of disorders, treatments, treatment availability, and the ability to communicate about health issues with others (Berkman, Davis, & McCormack, 2010; Kreps et al., 1998). In addition, mental health literacy is the ability to identify specific mental disorders, knowledge of mental disorder causes, and attitudes that facilitate recognition and help seeking, among other items (Jorm, 2000).
Poor mental health literacy is a barrier to treatment in part because people do not seek help for symptoms that they have not identified as disordered (Gulliver, Griffiths, & Christensen, 2010; Jorm, 2000; Thompson et al., 2008). Once the presence of a mental disorder is identified, however, people are more likely to seek treatment, and receive help when they do (Jorm, 2012). For example, young people are more likely to seek treatment from a general practitioner or mental health professional for PTSD and depression when they have been able to identify their symptoms as a particular psychiatric disease (Yap, Reavley, & Jorm, 2014). Also, older adolescents who believe they have a mental disorder and convey this to their general practitioner are more likely to receive a specific, correct diagnosis from them (Haller, Sanci, Sawyer, & Patton, 2009). Thus, people’s ability to identify their own symptoms as indicating a mental disorder is an important step in receiving treatment.\(^1\) Finally, recognition of disordered behaviors as a particular disorder is associated with less stigmatizing attitudes towards those disordered behaviors (Yap, Reavley, Mackinnon, & Jorm, 2013b), which may in turn increase treatment seeking.

Peer advice is also a key factor influencing whether a person with a mental health disorder seeks treatment, according to the Network Episode Model, a highly influential model mapping out the numerous and complex pathways to mental health treatment (Pescosolido, 1992; Pescosolido, 2006; Pescosolido and Boyer, 1999; Perry and Pescosolido, 2015). Lay people start to formulate ideas about what might be wrong with a family member or friend well before their peers ever see a clinician and are often the first to identify a problem and recommend help (Pescosolido, Brooks-Gardner, & Lubell, 1998). Further, people are more likely to seek

\(^1\) Of course, disorder identification is only a first step; people must also take subsequent action to receive treatment benefits (Jorm, 2000).
professional help for a disorder when a peer explicitly recommends that they get treatment for it (e.g., Dew et al., 1991). Thus, peer identification of mental disorders plays a key role in helping those with symptoms receive treatment, and is as important as self-identification of mental disorders.

Mental health literacy includes, in a sense, the ability to identify DSM disorders. In a number of countries, including the U.S. (in which the current study was conducted), mental disorders are named and defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM, 5th ed. [DSM-5]. American Psychological Association [APA], 2013), making the DSM of central importance for people’s knowledge of mental disorders (Kutchins & Kirk, 1997). Also, anyone in the U.S. with mental disorder symptoms who receives treatment must have an official DSM disorder diagnosis for insurance reimbursement.

Finally, health literacy is best understood in appropriate context. That is, the extent of health literacy that is most likely to benefit an individual varies based on that person’s age, social and community setting, past health issues, and personal experiences (Nutbeam, 2008). Given the importance of health literacy for treatment seeking, and the important role of context in health literacy (Nutbeam, 2008), we turned to the literature on people’s concepts of mental disorders to better understand how context affects mental disorder assessments.

**College Students’ Mental Disorder Concepts and the Influence of Context**

Research in cognitive psychology suggests that college students’ ability to identify mental disorders is influenced by their knowledge of mental disorders, or their mental disorder concepts. Specifically, there is evidence that college students conceptualize mental disorders as sets of symptoms that participate in cause and effect relationships with one another (Kim & Ahn, 2002). Their judgments appear to reflect the belief that symptoms stem from a central cause,
which gives each mental disorder its unique symptoms (Ahn, Flanagan, Marsh, & Sanislow, 2006). For example, a traumatic event cause may lead to behavioral symptoms such as avoidance of trauma reminders.

Perhaps because college students tend to represent mental disorders as causal theories (Kim & Ahn, 2002), they also tend to use causal context to understand and identify them. One example of this is that college students rely on the initial causal link in judging what type of treatment would best treat a hypothetical disorder (Yopchick & Kim, 2009). Classic work in causal reasoning has shown that viewing a connected set of information in a chain is one way to draw causal connections between those elements (Einhorn & Hogarth, 1986). For example, disordered symptoms may be viewed as stemming from a traumatic life event, which in turn leads to an intense reaction, which then causes downstream behavioral symptoms, such as avoiding reminders of the trauma.

People are prone to search for causal explanations for odd behaviors in others (Hastie, Schroeder & Weber, 1990). Once they find such an explanation, it influences the way they view those odd behaviors (Ahn, Novick, & Kim, 2003). There is a growing body of evidence demonstrating an understanding-normality effect in lay assessment, wherein a plausible, causal, life-event explanation (e.g., a severe car accident) for various disordered behaviors gives college students the intuitive sense of understanding those disordered behaviors (Ahn et al., 2003). This perception of understanding triggers the inference that the bizarre behaviors are less abnormal, less in need of treatment, and merit less social avoidance than the same behaviors presented without a life-event explanation, even though those behaviors continue to cause the same significant levels of distress and disability as they would if not explained (Kim & LoSavio, 2009; Ahn et al., 2003; Weine & Kim, 2015).
Even clinicians explain away mental disorder symptoms when they learn of a proportionately negative life-event that caused them, contrary to DSM specifications that diagnoses should be made independently of context. In work demonstrating a proportionate-response effect, clinicians judged cases of life events and subsequent behaviors that were proportionate in strength and valence (e.g., negative life events and disordered behaviors) to be less psychologically abnormal and less difficult to understand than cases of life events and subsequent behaviors that were disproportionate in strength and valence (e.g., negative life events and non-disordered behaviors; Kim, Paulus, Gonzalez, & Khalife, 2012a). In this work, we investigate whether college students also demonstrate such a proportionate-response effect. That is, we ask whether they find disordered behavioral symptoms to be less psychologically abnormal and less difficult to understand in proportionate, causal life-event context. In addition, we investigated whether college students can identify posttraumatic stress disorder (PTSD) and depression (major depressive disorder; MDD), as defined by the most recent versions of the DSM (DSM–IV–TR; APA, 2000; DSM-5, APA, 2013).

**Posttraumatic Stress Disorder and Major Depressive Disorder Identification in Context**

MDD is characterized by two weeks or more of sadness or low mood that persists almost all of the time and is accompanied by loss of pleasure in daily activities, weight changes, insomnia, fatigue, poor ability to concentrate, and other symptoms that interfere with major life activities. In the two recent versions of the DSM, there is no context required for a MDD diagnosis (as is the case for most disorders).

On the other hand, the identification of appropriate context is explicitly required for a PTSD diagnosis. The DSM-IV-TR (APA, 2000) indicated that a PTSD diagnosis is appropriate only if: The person was exposed to a traumatic event (actual or threatened death, serious injury,
or sexual violence; the A1 criterion), the person’s exposure to a traumatic event was accompanied by a reaction of intense fear, helplessness, or horror (the A2 criterion), and the person displayed PTSD behavioral symptoms (e.g., intrusive, unwanted, and distressing memories of an event, avoidance of reminders of an event, exaggerated startle response, etc.). Both the event and reaction criteria qualify as context because they provide background context for the observable behaviors that follow. In contrast, the current edition (DSM-5, APA, 2013) removed the A2 criterion from PTSD diagnosis, such that it includes only the identification of a traumatic event (the A1 criterion) and PTSD behavioral symptoms.

The removal of the A2 criterion from PTSD was controversial because recent survey and interview research demonstrated that an intense immediate reaction of fear, helplessness, or horror (A2 criterion), and not the experience of an objectively traumatic event (A1 criterion), predicted the onset of PTSD behavioral symptoms (Boals & Schuettler, 2009; Bodkin, Pope, Detke, & Hudson, 2007; Gold, Marx, Soler-Baillo, & Sloan, 2005; Rubin & Feeling, 2013). However, many people report non-A2 immediate reactions at the time of a traumatic event (e.g., anger, shame, numbness, or a military-trained response; Adler, Wright, Bliese, Eckford, & Hoge, 2008; Hathaway et al., 2010), and as such were not eligible to receive a DSM-IV-TR (2000) PTSD diagnosis, even if all other symptoms were met. Ultimately, the DSM-5 (APA, 2013) decided to give greater weight to the latter research, and removed A2 from the PTSD diagnostic criteria.

College students are exposed to DSM disorders through multiple channels. From the DSM-III (APA, 1980) up through the current day, DSM disorder descriptions and definitions make their way into public knowledge via newspaper articles, books, and radio programs (Lees-Haley & Dunn, 1994; Huon, Brown, & Morris, 1988), prototypical film and television depictions
(Hyler, 1988), television talk shows (Rasmussen & Ewoldson, 2013), and websites (one of the major sources from which older adolescents seek mental health information; Kauer, Mangan, & Sanci, 2014). In addition to absorbing knowledge of DSM definitions of mental disorders and examples of disordered people through these media, college students are likely to map their mental disorder concepts onto those used by professionals (Shaw, 2002). In order to facilitate disorder identification, it may be beneficial in practice for college students to understand each disorder, including PTSD and MDD, as clinicians are likely to understand it (that is, to the extent that clinicians adhere to the DSM; see Whooley, 2010), (Haller et al., 2009).

PTSD is a mental health issue affecting many college students and campuses. One fifth of college students have PTSD or sub-threshold PTSD symptoms (20%; Smyth, Hockemeyer, Heron, Wonderlich, & Pennebaker, 2008). According to the DSM-5 (APA, 2013), PTSD is triggered by the experience of a traumatic event, which 84% of college students have experienced in their lifetime (Vrana & Lauterbach, 1994). Furthermore, some types of traumatic events may be targeted at college students specifically. For example, college campus shootings have increased in recent years (Kay, 2010), and exposure to a campus shooting leads a subset of exposed college students to develop PTSD (Littleton, Kumpula, & Orcutt, 2011; Hughes, 2011).

Depression is also a major mental health problem for colleges. One half of college students self-report one or more depression symptoms (53%; Furr, Westefeld, McConnell, & Jenkins, 2008), and thus could benefit from symptom alleviation. Further, major depressive disorder (MDD) is the most common mood disorder experienced by college students, with an estimated 7% of college students qualifying for an MDD diagnosis in a 12-month period (Blanco et al., 2008). The rate of college students seeking alleviation from depression symptoms is not only high (Furr et al., 2008), but actually increased markedly between 2000 and 2007 (Kay,
2010). Given the prevalence of PTSD and MDD on college campuses, college students are very likely to know about PTSD and major depression from personal or secondhand experience, or word-of-mouth.

Our primary question was whether life-event context influences college students’ ability to identify mental disorders, specifically PTSD and MDD. In particular, we were interested in these two mental disorders because they are relevant to college students: They are likely to have these disorders, wish to alleviate their symptoms, and are likely to know about them first- or second-hand. It is important for college students to be able to identify PTSD and MDD, to increase their chances, and their peers’ chances, of receiving treatment. Further, it was unclear whether life-event context helps college students explain away the presence of these disorders (as in Ahn et al., 2003), particularly PTSD, which includes life-event context in its diagnostic criteria.

Past work using vignettes describing hypothetical people with disordered behavioral symptoms showed that members of the general public were able to identify clear-cut cases of PTSD in a vignette about 67% of the time (Merritt, Tharp, & Furnham, 2014), but that older adolescents were able to do so only about 34% of the time (Reavley & Jorm, 2011a). Members of the general public, older adolescents, and college students alike were able to identify a clinically unambiguous case of depression in a vignette about 74% of the time (Reavley & Jorm, 2011a; Reavley & Jorm, 2011b; Reavley, McCann, & Jorm, 2012). Thus, although people’s ability to identify clear-cut cases of PTSD and depression has been investigated, whether college students are influenced by a range of life-event contexts and reaction contexts in identifying these disorders, and are thus susceptible to explaining away the presence of these disorders, remains to be directly measured.
Overview of Current Study

In the current study, college students assessed hypothetical individuals described in short vignettes, following similar past work (Link et al., 1999; Pescosolido et al., 2010; Reavley & Jorm, 2011a). The vignette study approach allowed us to carefully control the context presented with each set of behavioral symptoms. Each vignette described behaviors in the context of events that were either traumatic or mildly stressful, as well as reactions to those events that were either intense or mild. The behaviors in each vignette described PTSD symptoms, MDD symptoms, or subclinical, mildly distressed behaviors.

We collected college students’ open-ended disorder identifications for each vignette. This task was intended to map as closely as possible to the real world situation in which students consider whether a peer has a particular disorder. Although past work often measured people’s ability to identify disorders by asking what was wrong with the person (e.g., Yap, et al., 2013b), we wanted to focus on their identification of specific disorders. The literatures on the understanding-normality effect (e.g., Ahn et al., 2003; Kim & LoSavio, 2009) and the proportionate-response effect (Kim et al., 2012a) suggest that people downplay disordered symptoms if they feel they can understand the origin of those symptoms. Thus, we wanted to ask for disorder identification using a open-ended dependent variable that, by simply asking whether the participant could identify any disorders in the vignette described, would allow us to tease apart people’s ability to detect mental disorders from their opinions about whether behavioral symptoms represent a problem. We also measured judgments of psychological abnormality and difficulty understanding the behaviors to capture whether the proportionate-response effect occurs during student assessments of disordered behaviors in this study.

Research Questions and Alternative Hypotheses
**Question 1a: Are college students influenced by context in identifying mental disorders?** Given past work on the *understanding-normality effect* showing that proportionate life-event context influences assessments of disordered behaviors in clinicians (Kim et al., 2012a), we asked whether proportionate context *increases or decreases* how often college students identify PTSD and MDD in accord with current (i.e., *DSM-IV-TR*; APA, 2000; *DSM-5*; APA, 2013) standards. One hypothesis is that college students will identify both disorders less often given any proportionate context because the context helps them explain away the presence of a disorder (i.e., the *proportionate-response effect*; Kim et al., 2012a).

An alternative is that college students will identify both disorders more often in proportionate context. For example, for PTSD behaviors, people identify PTSD more often given a traumatic context because they are under the impression that it is part of PTSD itself, as the *DSM-IV-TR* and *DSM-5* diagnostic criteria suggest (APA, 2000; APA, 2013). For MDD behaviors, people may identify MDD more often in proportionate context if they hold the belief that MDD behaviors are typically preceded by a traumatic life-event or intense reaction, and are not as likely to appear at random.

Another hypothesis is that college students will identify PTSD and MDD differently in context, as suggested by the *DSM-IV-TR* and *DSM-5* diagnostic criteria (APA, 2000; APA, 2013). That is, they may plausibly identify PTSD more often in proportionate context, but not be influenced by context when identifying MDD.

Yet another alternative is that proportionate context will not affect college students’ identification of either disorder. This would demonstrate that the presence of proportionate context does not help students identify PTSD or MDD, and that only the presence of behavioral symptoms helps college students identify these disorders. For PTSD, the latter finding would not
be in line with any recent prescriptive models of diagnosis, but for MDD, it would align with the DSM-IV-TR (APA, 2000) and DSM-5 (APA, 2013) diagnostic criteria.

**Question 1b: Specifically, is college students’ ability to identify mental disorders systematically influenced only by traumatic events, only intense reactions, only the presence of both traumatic events and intense reactions, or none of these contexts?** If the (DSM-5, APA, 2013) is taken to be the most current, generally accepted prescriptive model for disorder identification, then PTSD ought to be identified in traumatic life event context, regardless of the type of reaction context, whereas MDD ought to be diagnosed without necessarily being influenced by any type of context. That is, one hypothesis is that college students’ reliance on life-event context is dependent on the disorder in question.

On the other hand, it is possible that college students are influenced by proportionate life-event or reaction contexts in other ways when identifying these disorders. An alternative hypothesis is that traumatic event context, but not intense reaction context, will influence college students’ identification of both disorders. College students may simplify judgments by following a heuristic in which behaviors are understood best in the context of an initial causal link (the event) rather than a middle causal link (the reaction). As mentioned earlier, college students rely on the initial causal link in judging what type of treatment would best treat a hypothetical disorder (Yopchick & Kim, 2009).

Yet another alternative hypothesis is that intense reaction context, but not traumatic event context, will influence students’ ability to identify these disorders, as the reactions most immediately temporally precede the symptoms and could be perceived as most clearly responsible for bringing them about (Einhorn & Hogarth, 1986). For example, in the case of PTSD behaviors, people may make judgments in alignment with the body of work demonstrating
a strong link between intense reactions and the onset of PTSD symptoms (e.g., Boals & Schuettler, 2009).

A third alternative hypothesis is that only the presence of *both* a traumatic event and an intense reaction will influence students’ ability to identify these disorders. Past work by Einhorn & Hogarth (1986) suggests that people feel that they best understand events, reactions, and symptoms when all are proportionate in strength and negativity. If so, then people might be able to identify disorders only in the context of a proportionately strong and negative traumatic event and intense reaction, the three of which conjointly point to the need for diagnosis.

Finally, it is possible that none of these contexts will influence students’ disorder identifications. For example, people may believe the sole initial cause of mental disorders is their underlying biological basis (Ahn et al., 2006), and thus will not be influenced by the life event and reaction contexts manipulated in the current study.

**Methods**

The Northeastern University Internal Review Board approved all experiment protocols.

**Participants**

We recruited 96 undergraduate college students (64 female) at Northeastern University with no clinical experience, who participated for partial course credit. Eighty participated in the main study and 16 participated in the manipulation check. Participants reported a mean age of 19 years (range: 18-23), and self-identified as White (65%), Asian (19%), and Black (4%), or multiple races (8%; the remaining participants declined to report race). In addition, 10% self-identified as Hispanic.

**Materials**
Participants judged a total of six vignettes, each of which described a hypothetical person in two paragraphs, which were four to six sentences each. The first paragraph labeled “Recent Experiences” included a description of an event the person experienced (e.g., “they were pinned down by sniper fire”), and a description of the person’s reaction to that event (e.g., “she became completely paralyzed with fear”), roughly equated to each other for length. The event and reaction were both included in the same paragraph to meet the DSM-IV-TR (APA, 2000) intense reaction criterion (e.g., to clearly specify that the intense reaction was in response to the traumatic event). In each vignette, we wrote reaction text tailored to each event, and presented it such that each event was coherently described as being immediately followed by the person’s accompanying reaction. Please see Table A1 for a partial example vignette, and Appendix A for full vignettes.

The second paragraph of each vignette, labeled “Subsequent Behaviors” included five behavioral symptoms that described PTSD (taken from Criteria B-H from the DSM-IV-TR and DSM-5; e.g., “…she felt plagued by this event.”), or five behaviors that described Major Depressive Disorder (MDD; e.g., “She felt extremely lethargic…”), or five behaviors that did not meet diagnostic criteria for any disorder (distressed behaviors; e.g., “She felt uneasy but went on with her everyday life.”).

We asked people to reason about MDD behaviors as a comparison disorder to PTSD because neither event nor reaction context is required for a diagnosis (DSM-5, APA, 2013). Distressed behaviors described mild, non-disordered behaviors, which are an appropriate control to both sets of disordered behavioral symptoms because they describe someone coping with a traumatic life event in a way that does not indicate a disorder (Schwartz & Link, 1989). Each behavioral vignette described behaviors that began or continued three months or more after the
“Recent Past” to meet *DSM-IV-TR* (APA, 2000) and *DSM-5* (APA, 2013) criteria that PTSD symptoms begin at least one month after the event, and to differentiate behavioral vignettes from Adjustment Disorder and Acute Stress Disorder, in which symptoms may begin and end within one month (Acute Stress Disorder) or begin within zero to three months (Adjustment Disorder) after an event (APA, 2013).

**Event and reaction criteria.** The topics of the events described in the vignettes were either modified from Kim et al. (2012a), drawn from the *DSM-IV-TR* casebook (Spitzer, Gibbon, Skodol, Williams, & First, 1994), or drawn from published studies of people’s experiences leading to PTSD (Matsakis, 1996; Rothschild, 2000). Each event introduced the person briefly, then described a discrete occurrence including the person’s whereabouts, experiences, actions, and a resolution (e.g., “When it was over...”) to indicate that the event was a single occurrence, now concluded. See Table A1 for a partial sample vignette. The events were either traumatic or non-traumatic (everyday events). *Traumatic events* were defined by strict adherence to the A1 criterion from both the *DSM-IV-TR* and *DSM-5*, and described an objectively traumatic event involving actual or threatened death, serious injury, or sexual violence, which was experienced by a person directly or indirectly, or happened to a person’s family member or friend (e.g. receiving an injury in a severe car accident; *DSM-IV-TR*, APA, 2000; *DSM-5*, APA, 2013). *Everyday events* were designed to fall markedly short of meeting either the *DSM-IV-TR* (APA, 2000) or *DSM-5* (APA, 2013) PTSD A1 criterion for a traumatic event (e.g., did not describe an objectively traumatic event). Four of the everyday events described a less severe form of its corresponding traumatic event vignette (e.g. experiencing a fender-bender); two of the everyday events described a secondhand experience of its corresponding traumatic event vignette (e.g. reading about a bad car accident in the newspaper).
The reactions were described as being either intense or mild. *Intense reactions* were defined by strict adherence to the A2 criterion in *DSM-IV-TR* (APA, 2000), as a severe internal reaction of fear, helplessness, or horror that encompassed someone’s thoughts, feelings, and emotions at the time of an event (e.g., feeling afraid and thinking about fleeing; *DSM-IV-TR*, APA, 2000). *Mild reactions* did not meet the *DSM-IV-TR* A2 criterion (e.g., were not an intense reaction of fear), and described someone’s mildly negative or neutral thoughts and feelings at the time of an event (e.g., feelings of surprise and thoughts of concern; APA, 2000). Again, please see Table A1 for an example. Reactions never described a potential behavioral symptom of one of the disorders. For example, the reaction could not be described as someone feeling jumpy, because this could potentially be interpreted as indicating the presence of the PTSD criterion E, a heightened startle reflex (*DSM-5*, APA, 2013).

**Manipulation Check**

Sixteen undergraduate psychology students (10 female) independently rated each event for how traumatic it seemed (on a Likert scale of 1-9, where 1=not at all traumatic and 9=extremely traumatic), and each reaction for how severe it seemed (on a Likert scale of 1-9, where 1=not at all severe and 9=extremely severe), in randomized order. Paired samples t-tests revealed that participants found traumatic events (*M* = 7.52, *SD* = .77) to be more traumatic than everyday events (*M* = 3.67, *SD* = 1.34; *t*(15) = 9.15, *p* < .001), and intense reactions (*M* = 7.17, *SD* = 1.04) to be more severe than mild reactions (*M* = 3.55, *SD* = .96; *t*(15) = 12.14, *p* < .001). Further, all means ran in this same direction for each vignette.

**Procedure**

Participants read a total of six vignettes. The between-subjects factors were event and reaction condition. Each participant read about one of the two event types (traumatic or mild)
combined with one of the two reaction types (intense or mild). The within-subjects factor was behavior type. Two vignettes described PTSD behavioral symptoms, two described MDD behavioral symptoms, and two described distressed behaviors. For each vignette, people were asked to type in up to three disorders, and were explicitly instructed that one was required, which could be “none”. They were instructed to rate the likelihood of the presence of each disorder on a scale of 1-9, where 1= very unlikely and 9=very likely (please see Table A2 for exact text of all dependent variables). They were also asked to make judgments of psychological abnormality (on a scale of 1-9, where 1= very psychologically normal and 9=very psychologically abnormal), and judgments of difficulty understanding the subsequent behaviors (on a scale of 1-9, where 1= very easy to understand and 9=very difficult to understand; again, please see Table A2).

Participants answered the same question for all six vignettes before moving on to answer the second question for all six vignettes. This was done so that the first dependent variable measured could not be contaminated by subsequent dependent variables. For example, this was done so that we could be sure that open-ended diagnoses did not differ when they were made first compared to when they were made after psychological abnormality judgments. The vignette was always reprinted in full on the page with each dependent variable, such that participants did not have to keep the vignette details in mind between question blocks. The order of dependent variables was randomized, and the order of vignettes was also randomized for each question.

**Results and Discussion**

**Primary Analyses: Open-Ended Disorder Identifications in Context**

We asked whether proportionate context influences disorder identifications, and if so, exactly what kinds of context (i.e., events, reactions, or both in conjunction). We scored participants as identifying a disorder when they provided the same disorder response for both
vignettes of the same behavior type (e.g., indicated PTSD for both PTSD vignettes). We next examined the effects of behavior type, event context, and reaction context on the proportion of open-ended disorder identifications of primary theoretical interest: identifications of PTSD, identifications of depression or MDD, and indications that no disorder was present (please see Appendix B for additional analyses that are not related to the main questions of interest). To do so, we ran three separate 3 (Behavior Type: PTSD, MDD, distressed) x 2 (Event Type: Traumatic Event, Everyday Event) x 2 (Reaction Type: Intense Reaction, Mild Reaction) repeated-measures ANOVAs, one each on the proportion of identifications of PTSD, MDD, and “none.” Also, to examine the likelihood judgments provided by participants with each disorder identification, we ran three separate 2 (Event Type: Traumatic Event, Everyday Event) x 2 (Reaction Type: Intense Reaction, Mild Reaction) ANOVAs on the likelihood ratings college students gave to their own disorder identifications of our primary interest: Identifications of PTSD for PTSD behaviors, identifications of “depression” or MDD for MDD behaviors, and indications that no disorder was present for distressed behaviors.

**PTSD identification in context.** The 3 (Behavior Type: PTSD, MDD, distressed) x 2 (Event Type: Traumatic Event, Everyday Event) x 2 (Reaction Type: Intense Reaction, Mild Reaction) repeated-measures ANOVA performed on the proportion of PTSD identifications revealed a main effect of behavior type and a main effect of event type that were both qualified by an interaction of behavior type and event type (F[2,152] = 7.31, p = .001, ηp² = .09). Please see Figure A3. There were no other main effects or interactions (all p ≥ .371, all ηp² ≤ .011).

To examine the interaction, six Bonferroni-corrected paired samples t-tests were run, three in traumatic event context only, three in everyday event context only. In traumatic event context only, PTSD behaviors were identified as PTSD more often than either MDD behaviors...
Further, in traumatic event context only, MDD behaviors were identified as PTSD more often than distressed behaviors ($t[39] = 2.88, p = .006$). Next, in everyday event context only, PTSD behaviors were identified as PTSD no more often than MDD behaviors ($t[39] = 1.36, p = .183$) or distressed behaviors ($t[39] = 1.36, p = .183$). There was also no difference between MDD and distressed behaviors ($t[39] = 0.00, p = 1.000$). Again, please see Figure A1. Thus, college students identified PTSD *most often* for PTSD behaviors when given traumatic life-event context and second most often for MDD behaviors when given traumatic life-event context.

Next, three independent samples t-tests revealed that PTSD behaviors were identified as PTSD more often in traumatic event context than in everyday event context ($t[78] = 4.56, p < .001$). Also MDD behaviors were identified as PTSD more often in traumatic event context than in everyday event context ($t[78] = 3.30, p = .001$). However, distressed behaviors were identified as PTSD no more often in either event context ($t[78] = 1.38, p = .170$). Thus, traumatic event context was sufficient to motivate a PTSD diagnosis for PTSD and MDD behaviors, but not distressed behaviors. Also, traumatic event context motivated a PTSD diagnosis less so for MDD behaviors than for PTSD behaviors.

Overall, the main analysis revealed that participants found traumatic event context for disordered behaviors to be likely indication that PTSD was present. This effect was increased when the disordered behaviors described PTSD behavioral symptoms.

Finally, we also ran a 2 (event context) x 2 (reaction context) ANOVA on likelihood judgments participants assigned to their own identifications of PTSD for PTSD behaviors. These analyses revealed that college students also rated their own indications of PTSD for PTSD behaviors to be more likely in traumatic event context ($M = 7.31, SD = 1.74$) than in everyday
event context ($M=5.25$, $SD=.87$; $F[1,52] = 16.27$, $p < .001$, $\eta^2_p = .25$). Please see Figure A4.

There was neither a main effect of reaction type nor an interaction (all $ps >= .133$). Thus, not only did college students identify PTSD more often in traumatic life-event context than in everyday life-event context, they also found their own judgments to be more likely.

Overall, college students were not swayed by reaction context at all when identifying PTSD, as there were no main effect or interactions involving reaction context. PTSD identifications seemed to take into account only the context required for a PTSD diagnosis in the current DSM-5 (i.e., only traumatic event context; APA, 2013). Further, this context increased students’ judgments of the likelihood of their own identifications. Generally, student judgments did not align with past DSM-IV-TR (APA, 2000) diagnostic criteria for PTSD (i.e., requiring the conjunction of a traumatic event and intense reaction), or with research that indicates that intense reactions, not objectively traumatic events, precede PTSD symptoms (e.g., a lockdown drill might not be judged traumatic by most, but it might lead to PTSD symptoms if accompanied by an intense reaction of fear; Boals & Scheuttler, 2009). Please see Difficulty understanding judgments discussion (p. 30), Summary of Findings, and Implications for Mental Health Literacy (p. 31), and (p. 32) for further discussion of these results.

**Depression identification in context.** The 3 x 2 x 2 repeated measures ANOVA performed on depression or MDD identifications revealed only a main effect of behavior type ($F[2,152] = 108.90$, $p < .001$, $\eta^2_p = .589$). There were no other main effects or any interactions (all $p >= .113$, all $\eta^2_p < .028$). Bonferroni-corrected paired-samples t-tests comparing depression identifications for PTSD, MDD, and distressed behaviors revealed that MDD behaviors were identified as depression more often than either PTSD behaviors ($t[79] = 10.01$, $p < .001$) or distressed behaviors ($t[79] = 11.79$, $p < .001$). Further, PTSD behaviors were identified as
depression no more often than distressed behaviors ($t[79] = 2.04, p = .045$). Please see Figure A2). That is, MDD was identified as depression regardless of event or reaction context, and only for MDD behaviors. Thus, participants did distinguish between disordered behavior types when identifying depression.

College students identified depression for MDD behaviors regardless of context, in line with recent and current nosologies ($DSM-IV-TR$, APA, 2000; $DSM-5$, APA, 2013). To be clear, it is not the case that the $DSM-IV-TR$ (APA, 2000) and $DSM-5$ (APA, 2013) forbid depression diagnoses in context (but please see Footnote 2); rather, they simply omit context within the diagnostic criteria, just as for the majority of their listed disorders. One of our hypotheses was that context that is proportionate to MDD behaviors might help explain away the presence of a disorder via an extension of the proportionate-response effect, by helping the disordered behaviors seem less psychologically abnormal and perhaps less likely to be a disorder. For college students, it did not do so, perhaps because they are so familiar with the symptoms of depression that they are able to identify it without getting distracted by the context, though this interpretation is entirely speculative and needs further study. Indeed, about half of college students self-report depression (Furr et al., 2008) so college students very likely have had the opportunity to either experience depression or view peers’ depressive behaviors in a wide variety of contexts.

The 2x2 ANOVA on likelihood judgments did not reveal a main effect of event or reaction type (all $ps > .181$), but there was an interaction ($F[1,66] = 5.69, p = .020, \eta^2_p = .08$). Bonferroni-corrected independent samples t-tests were conducted to examine the interaction. In traumatic event context, the two reaction contexts did not affect judgments differently ($p = .219$). However, in everyday event context, MDD behaviors were judged less likely to be depression or
MDD after a mild reaction ($M = 6.56$, $SD = 1.52$) than after a severe reaction ($M = 7.63$, $SD = 1.32$; $t[30] = 2.12$, $p = .042$). Furthermore, in severe reaction context, the two event types did not affect judgments differently ($p = .490$). In mild reaction context, however, MDD behaviors were judged less likely to be depression or MDD after an everyday event ($M = 6.56$, $SD = 1.53$) than after a traumatic event ($M = 7.90$, $SD = 1.14$; $t[30] = 2.80$, $p = .009$). Please see Figure A4. Generally, people rated their own indications of depression or MDD for MDD behaviors to be less likely in the context of an everyday event paired with a mild reaction compared to all three other event and reaction combinations (e.g., a traumatic event paired with a severe reaction).

That is, college students rated depression to be less likely in everyday event / mild reaction context compared to the other three event reaction combinations (i.e., traumatic event / intense reaction, traumatic event / mild reaction, everyday event / intense reaction), even though this was not reflected in the proportion of disorder identification judgments. One possible reason for this is that college students have some notion that depression is typically precipitated by something, rather than occurring without an identifiable trigger. Thus, even though college students were not influenced by context when identifying depression for MDD behaviors, they rated such identifications less highly when a causal precipitant (i.e., a traumatic event, intense reaction, or both) was not easily identified.

“No disorder” identification in context. The 3 x 2 x 2 repeated measures ANOVA performed on “none” identifications revealed only a main effect of behavior type ($F[2,152] = 75.29$, $p < .001$, $\eta^2_p = .500$). There were no other main effects or any interactions (all $p \geq .207$, all $\eta^2_p < .021$). Bonferroni-corrected paired-samples t-tests revealed that distressed behaviors were given a “none” response more often than either PTSD behaviors ($t[79] = 10.01$, $p < .001$) or MDD behaviors ($t[79] = 11.79$, $p < .001$). Further, PTSD behaviors were not given a “none”
indication any more often than MDD behaviors ($t[79] = 2.04, p = .045$). Please see Figure A3. That is, distressed behaviors were identified as not describing any disorder regardless of event or reaction context, but PTSD and MDD behaviors were not judged to be good candidates for such a response (again please see Figure A3).

The 2x2 ANOVA on “none” likelihood judgments for distressed behaviors did not reveal a main effect of event type, reaction type, or an interaction (all $ps > .137$). Please see Figure A4.

**Other disorders identified.** Not all disorder identifications constituted the appropriate response for each vignette with respect to the *DSM-IV-TR* (APA, 2000) or *DSM-5* (APA, 2013; please see Tables A3 and A4). Given space to provide up to three disorder responses, college students indicated a wide range of *DSM* disorders. In total, students suggested 20 different *DSM* disorders (please see Table A3 for response counts by behavior type, and Appendix B for a more complete description of the results of the open-ended disorder identification task that were not of main theoretical interest.) In addition to PTSD and MDD identifications examined in the main analyses, college students indicated the superordinate category of anxiety disorder a total of 82 times across all 480 vignettes (17%, please see Table A3). Thus PTSD, depression, anxiety, and indications that no disorder was present were the most common responses on our open-ended measure across the three behavior types. Exactly why this might be is a matter of speculation at present, but it is possible that college students may have identified these particular disorders most often because they have been encountered frequently and are therefore highly available in memory. In line with this possibility, college students report wanting to alleviate symptoms of depression and anxiety disorders more than any other mental health disorders (Kay, 2010).

Interestingly, only 7% of all responses were neither a *DSM-IV-TR* (APA, 2000) or *DSM-5* (APA, 2013) disorder nor a “none” response (please see Tables A3 and A4 for details by
vignette and Appendix B for all responses). In total, college students made 13 trauma-related responses (e.g., “survivor’s guilt”), 8 that resembled a DSM disorder, proposed DSM disorder, or symptom from a DSM disorder other than a trauma-related disorder (e.g., “compulsive buying disorder;” “emotional detachment”), 9 that did not approximate a known diagnosis (e.g., “confident disorder;” “hyper empathy”), 14 states (e.g., “stress;” “nervousness”), and 3 acknowledging a lack of knowledge (e.g., “I don’t know”). It is not necessarily the case that these relatively rare non-DSM responses should be viewed as incorrect; for example, college students may actually have been trying to describe a DSM disorder that they were familiar with, but were unable to retrieve the precise terminology at the time of the study. Of course, they may simply have done their best to indicate a condition that they felt the person in the vignette possessed, using a label they found descriptive, and were not attempting to remember the name of a DSM disorder. Follow-up work will be needed to tease these possibilities apart.

Secondary Analyses: Judgments of Psychological Abnormality and Difficulty Understanding

For the secondary judgments of psychological abnormality and difficulty understanding, we ran 2 (Event Type) x 2 (Reaction Type) ANOVAs, conducted at the $\alpha = .05$ level, for each behavior type. Paired comparisons were Bonferroni-corrected. In all analyses, ratings for individual vignettes of the same behavior type were collapsed.

**Psychological abnormality judgments.** There was a main effect of event type for the following judgments: psychological abnormality judgments for PTSD behaviors ($F[1,76] = 9.09$, $p = .003, \eta^2_p = .11$) and MDD behaviors ($F[1,76] = 18.31, p < .001, \eta^2_p = .19$). Please see Figure A5. There were no main effects of reaction type (all $ps \geq .729$, all $\eta^2_p \leq .002$), nor any interactions (all $ps \geq .646$, all $\eta^2_p \leq .003$). Thus, both disordered behaviors were rated as less
psychologically abnormal in the context of a traumatic event than in the context of an everyday event. These findings were in concert with past work demonstrating the proportionate-response effect (Kim et al., 2012a), wherein proportionate life-event context made MDD and PTSD behaviors seem less psychologically abnormal to practicing clinicians. The current findings replicated this past work in judgments made by college students. For distressed behaviors, there were no main effects or interactions (all $ps > .227$; all $\eta_p^2 < .02$). That is, neither event nor reaction contexts influenced how psychologically abnormal the distressed behaviors seemed. It appears that college students find merely distressed behaviors to be equivalently psychologically abnormal regardless of the context in which they occur.

**Difficulty understanding judgments.** There was also a main effect for difficulty understanding judgments for PTSD behaviors ($F[1,76] = 53.15$, $p < .001$, $\eta_p^2 = .41$) and MDD behaviors ($F[1,76] = 58.62$, $p < .001$, $\eta_p^2 = .44$; see Figure A6). There were no main effects of reaction type (all $ps > .424$, all $\eta_p^2 < .01$), nor any interactions (all $ps > .807$, all $\eta_p^2 < .01$). Also, for distressed behaviors, there were no main effects or interactions (all $ps > .153$; all $\eta_p^2 < .03$). In line with our findings for psychological abnormality judgments, the disordered behaviors were rated as less difficult to understand in the context of a traumatic event than in the context of an everyday event, replicating past work demonstrating the proportionate-response effect with college students. Also, context did not affect how difficult it was for students to understand distressed behaviors.

One of our hypotheses was that context proportionate to PTSD and MDD behaviors might help explain away the presence of a disorder via the proportionate-response effect, by helping the disordered behaviors seem less psychologically abnormal and perhaps less likely to be an identifiable disorder. Instead, our data showed that proportionate event context appeared to
help people identify PTSD and did not affect MDD identifications at all. Thus, college students’
disorder identifications and judgments of psychological abnormality were affected in opposite
directions by event context, not supporting an extension of the proportionate-response effect to
disorder identifications. This suggests that lay people may not necessarily think that
psychological abnormality and the presence of a disorder go hand in hand. Indeed, our work
demonstrates that college students feel someone’s behaviors can be identified as disordered
regardless of whether they view those behaviors as psychologically abnormal to begin with. That
is, college students may hold the underlying belief that even if a peer’s behaviors are a
psychologically normal response to recent life events, and therefore are relatively easy to
understand, they still merit disorder identification because they fulfill certain known criteria.
More research is needed to determine why psychological abnormality judgments and disorder
identification judgments do not appear to be influenced by context in a similar way.

**Summary of Findings**

The issue that we addressed in this work is whether college students can identify mental
disorders in context. In this study, college students were able to identify these disorders in accord
with *DSM-5* (APA, 2013) diagnostic criteria (and also *DSM-IV-TR* [APA, 2000] criteria for
depression). The proportionate-response effect did not extend to the two open-ended disorder
identification responses of interest: PTSD and depression. That is, college students did not find
either PTSD or MDD behaviors any less indicative of a disorder given proportionate context than
without it. Rather, for both types of disordered behaviors, proportionate context did not decrease
the frequency with which they identified the relevant disorder. Of course, we do not claim that
college students have necessarily read about these disorders in the *DSM-IV-TR* (APA, 2000) or
*DSM-5* (APA, 2013) or explicitly learned about them some other way. Rather, we observe that
their judgments appear to align with these nosologies to a substantial degree, in accord with the most current nosology (*DSM-5*, APA, 2013).

Reaction context did not influence how college students thought about PTSD, MDD or distressed behaviors in this study; this was the case for diagnostic judgments, judgments of psychological abnormality, and even simple difficulty understanding judgments. Again, disproportionate reaction information did not make it *more difficult* for laypeople to understand any of the behaviors; instead, it simply had no effect. It could be that reaction information is not psychologically salient to college students, and/or not perceived as particularly important, and therefore they do not incorporate it into their concepts of other people’s behaviors.

Next, there was no evidence that college students relied *only* on the presence of PTSD behavioral symptoms to identify PTSD, another alternative hypothesis. Rather, it appears to be the case that college students feel that PTSD is the conjunction of any type of disordered behaviors and the presence of a traumatic event (although the only disordered behavioral symptoms tested in this study were PTSD and MDD). Although there is some overlap in the behavioral symptoms for PTSD and MDD (e.g., negative mood), each PTSD and MDD vignette was carefully written to describe only the desired disorder and no other. That is, MDD vignettes that described negative mood also described four other symptoms that qualified the hypothetical person for a MDD diagnosis, and insufficient symptoms to qualify them for any other diagnosis. Participants seemed somewhat sensitive to this, as they distinguished between the two somewhat, (diagnosing PTSD less often for MDD behaviors in traumatic event context than for PTSD behaviors). Future work is needed to understand whether *any* disordered behaviors plus a traumatic event motivates PTSD diagnoses in college students, or if this effect is specific to PTSD and MDD behaviors.
Implications for Mental Health Literacy

It does not appear to be the case that lack of ability to identify mental disorders is the only barrier to treatment for college students. That is, even though college students could identify disorders in context in this study, they still judged them to be less psychologically abnormal and less difficult to understand in such context. Past work with college students and lay people demonstrated that a hypothetical person’s alcohol use disorder behaviors were perceived as less in need of treatment in the context of a proportionately negative life-event, compared to the exact same alcohol use disorder behaviors presented without such context (Weine & Kim, 2015). We speculate that college students may yet be discouraged from seeking treatment or recommending it, even though a substantial number may be able to identify these disorders, because they believe the behaviors are an understandable response to negative life event. This possibility can be tested directly in future work.

Stigma may also contribute to students’ reluctance to seek treatment (Eisenberg, Downs, Golberstein, & Zivin, 2009). Stigma is a major barrier to treatment for PTSD and depression. Believing that either condition is indicative of personal weakness rather than illness decreases older adolescents’ treatment seeking from all professional sources of help (Yap, Reavley, & Jorm, 2013a). Further, greater stigma towards people with PTSD (i.e., desired social distance) correlates with lower acceptance of family support (Yap et al., 2013a). Yet, not all forms of stigma may be a deterrent to treatment. For example, war veterans are more likely to undergo treatment for depression when they perceive their symptoms as a health threat (Elwy et al., 2013), and older adolescents are more likely to seek treatment from a counselor for PTSD or depression when they believe their behaviors are a danger to themselves or others (Yap et al., 2013a). Thus, these particular negative beliefs about PTSD and depression might help frame
these disorders as a threat that needs to be combatted, rather than a personal dysfunction. Further work is needed to test whether reframing PTSD and depression as threats that can be defeated will ultimately lead to higher rates of treatment seeking among college students.

Our findings can inform future outreach interventions for college students’ health literacy. In this study, college students sometimes confused depression for PTSD, confused PTSD and depression with anxiety disorders, and suggested about 20 other *DSM-IV-TR* (APA, 2000) and *DSM-5* (APA, 2013) disorders all told. Future outreach programs to help college students better disentangle and clarify their concepts of these disorders could increase rates of accurate disorder identification amongst college students. Possible positive outcomes of such interventions include greater rates of treatment seeking (Yap, et al., 2014), increased communication effectiveness with health professionals (Haller et al., 2009), and decreased stigma towards mental disorders (Yap et al., 2013b).

**Limitations and Future Directions**

In this study, we employed carefully controlled vignettes and explicitly asked college students to identify the presence of specific disorders in the hypothetical people described in them. This allowed us to cleanly manipulate possible factors that might influence students’ ability to identify disorders, including the information about the disordered behaviors and background information about the person described. Though this method allowed us to maintain rigorous experimental control, it also necessarily differed from real world contexts in that students on college campuses may most often learn about peers’ behaviors via word of mouth or by direct observation and may not be directly asked to identify a disorder, although they are likely to observe, think about, and evaluate their peers’ behaviors on a regular basis. Whether
college students make similar judgments in more ecologically valid paradigms remains to be investigated.

Our sample consisted of college students at an urban university in the northeast U.S. Future work should expand the investigation of this aspect of mental health literacy – whether college students can identify mental disorders in context – to other college environments, as these may be very diverse (e.g., including suburban and rural campuses, as well as campuses in different regions and countries). Given the nature of the vignette study and the wealth of data provided by open-ended disorder identification responses, we limited our dependent variables to test one major aspect of mental health literacy: whether people can identify disorders (Jorm, 2012). We have not yet included many measures to gain a broad, multifaceted view of whether college students’ mental health literacy as a whole, across many disorders, is impacted by context. As Jorm (2000) pointed out, identification is only a first step, and is meaningless without further action that leads to treatment. This study is an initial investigation. Future work is needed to understand whether college students’ knowledge of appropriate treatment, treatment availability, and prevention strategies for a wide range of common disorders (Jorm, 2012), among other important aspects of mental health literacy, are influenced by context.
References


### Table A1

**Partial Sample Vignette – “Recent Past”**

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Reaction Type</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Traumatic Event</strong></td>
<td>Lucy teaches at an elementary school where there was a school shooting several months ago. She was with her class of children in the playground when they were pinned down by sniper fire. <em>Immobile and vulnerable she became completely paralyzed with fear. She felt like they were all going to die.</em></td>
</tr>
<tr>
<td><strong>Everyday Event</strong></td>
<td>Lucy teaches at an elementary school where there was a lockdown drill several months ago. She hadn’t known about the drill in advance because she had been out sick and missed the memo. <em>Immobile and vulnerable she became completely off guard and confused about what was going to happen. She tried to be confident for the children.</em></td>
</tr>
</tbody>
</table>

**Note:** This is a partial vignette for illustrative purposes. The actual vignettes described 4-5 times
more information (see Appendix A for complete text Reaction text, italicized here to differentiate it from event text, was not italicized in the main study.


### Table A2

**Dependent Variables**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Question Text</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open-Ended</td>
<td>Given [X]’s subsequent behaviors, what mental</td>
<td>Please respond to this question by typing the name of a mental disorder into</td>
</tr>
<tr>
<td>Diagnoses</td>
<td>disorder(s) is [X] likely to have, if any?</td>
<td>Diagnosis Box 1 below.(^a)</td>
</tr>
<tr>
<td>Psychological</td>
<td>How psychologically normal, or psychologically</td>
<td>Please choose your response on a scale of 1-9, where 1=very psychologically</td>
</tr>
<tr>
<td>Abnormality</td>
<td>normal, or psychologically abnormal, are [X]’s</td>
<td>abnormal and 9=very psychologically normal.</td>
</tr>
<tr>
<td></td>
<td>subsequent behaviors?</td>
<td></td>
</tr>
<tr>
<td>Difficulty</td>
<td>How easy or difficult is it to understand [X]’s</td>
<td>Please choose your response on a scale of 1-9, where 1=very easy to understand</td>
</tr>
<tr>
<td>Understanding</td>
<td>subsequent behaviors?</td>
<td>and 9=very difficult to understand.</td>
</tr>
</tbody>
</table>

\(^a\) Open-ended diagnoses instructions continued: If you believe no diagnosis is warranted, please type “none” into Diagnosis Box 1, and leave the other boxes blank. If there are any other diagnoses you think are likely, please type them into the subsequent Diagnosis Boxes 2 and 3 (one diagnosis per box). Then, please rate the likelihood of each diagnosis on a scale of 1-9, where 1= very unlikely and 9=very likely, by selecting a rating on the scale below each diagnosis box. If you believe no diagnosis is warranted, please also rate the likelihood of this answer on the scale of 1-9, where 1 = very likely and 9= very unlikely.
### Open-Ended Disorder Identification Responses by Behavior Type

<table>
<thead>
<tr>
<th>Open-Ended Diagnosis Responses</th>
<th>PTSD</th>
<th>MDD</th>
<th>Distressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTSD</td>
<td>77</td>
<td>46</td>
<td>19</td>
</tr>
<tr>
<td>Depression (including MDD)</td>
<td>42</td>
<td>117</td>
<td>15</td>
</tr>
<tr>
<td>&quot;None&quot;</td>
<td>7</td>
<td>10</td>
<td>103</td>
</tr>
<tr>
<td>Anxiety Disorder</td>
<td>34</td>
<td>21</td>
<td>27</td>
</tr>
<tr>
<td>Paranoia / Paranoid Personality Disorder</td>
<td>9</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Generalized Anxiety Disorder</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Specific Phobia</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Obsessive Compulsive Disorder</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Mood Disorder</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Panic Disorder</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bipolar Disorder</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Eating Disorder</td>
<td>0</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Attention Deficit Hyperactivity Disorder</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Insomnia Disorder</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Acute Stress Disorder</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Social Anxiety Disorder</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Disorder</td>
<td>Non DSM-IV-TR or DSM-5</td>
<td>2018</td>
<td>2019</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>------------------------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Alcohol Use Disorder</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Binge Eating Disorder</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Other Personality Disorder</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dissociative Disorder</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Borderline Personality Disorder</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Non DSM-IV-TR or DSM-5</td>
<td>23</td>
<td>11</td>
<td>13</td>
</tr>
</tbody>
</table>

**Disorder Responses**

| Total Responses | 231 | 224 | 189 |
Table A4

*Number (Percent of Total) of Open-Ended Disorder Identification Response Categories by Behavior Type*

<table>
<thead>
<tr>
<th>Behavior Type</th>
<th>Number of Vignettes</th>
<th>Total Responses</th>
<th>DSM Disorder Identifications</th>
<th>Non-DSM Disorder Identifications</th>
<th>&quot;None&quot; Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTSD</td>
<td>160</td>
<td>231</td>
<td>200 (87%)</td>
<td>23 (10%)</td>
<td>7 (3%)</td>
</tr>
<tr>
<td>MDD</td>
<td>160</td>
<td>224</td>
<td>203 (91%)</td>
<td>11 (5%)</td>
<td>10 (4%)</td>
</tr>
<tr>
<td>Distressed</td>
<td>160</td>
<td>189</td>
<td>73 (38%)</td>
<td>13 (7%)</td>
<td>103 (55%)</td>
</tr>
</tbody>
</table>

*Note.* Percentages in parentheses indicate the number of responses listed divided by the number of total responses.

*Total Responses* displays the total number of disorder identification responses college students provided for each behavior type. (This sum is larger than the number of vignettes because students were given space to identify up to three disorders for each vignette.)

*DSM Disorder Identifications* displays the number of *DSM-IV-TR* (APA, 2000) and *DSM-5* (APA, 2013) disorder identification responses for each behavior type.

*Non-DSM Disorder Identifications* displays the number of responses that were *not* a *DSM-IV-TR* (APA, 2000) disorder, *DSM-5* (APA, 2013) disorder, or “none” response.

*“None” Responses* displays the number of indications that no disorder was present.
Figure A1

Proportion of PTSD Identifications

Note: Error bars indicate standard errors.
Figure A2

Proportion of Depression Identifications

Note: Error bars indicate standard errors.
Figure A3

Proportion of “None” Identifications

Note: Error bars indicate standard errors.
Figure A4

Likelihood of Disorder Identifications by Event and Reaction Type

Note: Error bars indicate standard errors.
Figure A5

College Student Psychological Abnormality Judgments

Note: Error bars indicate standard errors.
Figure A6

College Student Difficulty Understanding Judgments

Note: Error bars indicate standard errors.
Interim Discussion

The previous manuscript, which was prepared for submission to a health communication journal, differs in a few important ways from the following manuscript, which was prepared for submission to a clinical psychology journal. These differences merit separate manuscripts prepared for publication. The health communication work sampled college students who had no significant clinical experience. On the other hand, the clinically oriented work (described next), sampled licensed, practicing, clinical psychologists and clinical trainees enrolled in a clinical psychology graduate program. The motivations for identifying PTSD differ between lay people and professionals. College students, clinical trainees and practicing clinicians may differ in what types of behaviors intuitively seem abnormal and therefore seem to merit a diagnosis given different circumstances. Also, there are differences between these populations in relative amounts of experience interacting with people who have disordered behaviors, relative amounts of experience thinking through diagnoses, and relative amounts of experience using the different versions of the DSM (e.g., DSM-IV-TR, APA, 2000 and DSM-5, APA, 2013).

Finally, work with these two populations informs different literatures. In the extensive literature on health literacy, which spans psychology, sociology, and health sciences, a major topic of interest is whether lay people can identify mental disorders at all (Jorm, 2000). We conducted the research reported in the previous manuscript because it was unknown whether college students can identify PTSD (and MDD) in context. Thus, our finding that college students can, indeed, identify PTSD in context a substantial amount of the time best informs the mental health literacy literature.

On the other hand, with regard to clinicians, disorder diagnosis is an integral part of clinical practice, and it would not be of interest simply to show whether or not professionals can
reliably identify PTSD, as this has already been accomplished in large-scale field research (Regier et al., 2013). The important and pressing issue currently at hand is that it is unknown exactly what types of context guide PTSD diagnoses by clinicians and clinical trainees, especially given the recently updated diagnostic criteria for PTSD diagnosis in the DSM-5 (APA, 2013). The DSM-5 (APA, 2013) had been introduced within a year of the current work, and during its development, considerable controversy surrounded their decisions regarding the exact types of context officially required for a PTSD diagnosis. Thus, the work undertaken with clinicians and clinical trainees best informs the clinical psychology literature, which is concerned with, in part, how clinicians diagnose disorders in practice. The results of our experiments with clinicians are important for informing our understanding of the clinical utility, or usefulness from clinicians’ perspective, of the current DSM-5 (APA, 2013). Thus, in the next manuscript, this research is presented in the context of this debate and these clinical issues.
Events, Reactions, and Behaviors: Clinical Assessment of Posttraumatic Stress Disorder

Erienne R. Weine, M.A. & Nancy S. Kim, Ph.D.

1 Department of Psychology, Northeastern University, 125 Nightingale Hall, 360 Huntington Avenue, Boston, MA 02115, U.S.A. (Location in which the work was performed)

* Corresponding author: Erienne R. Weine. Department of Psychology, Northeastern University, 125 Nightingale Hall, 360 Huntington Avenue, Boston, MA 02115, U.S.A., Tel: 520-603-3624; Fax: 617.373.8714; Email: e.weine@neu.edu.

We thank Dr. Dawn Cisewski and Dr. Susan Nissen for their professional clinical opinions on the wording of our vignettes and dependent variables. We thank Tara Armor, Sara Atlas, Marianna Ferris, Rachel Foulkes, Dunia Kassay, Britney Leibert, Allison Leibold, Jordan Nissen, Katharine Quain, Alicia Rogers, Angela Saulsbery, Dillon Smith, and Cara Vincelette for help with checking materials, programming, data collection, and coding.

Key Words: Posttraumatic Stress Disorder (PTSD), clinical assessment, traumatic events, reactions, cognitive processes
Abstract

There is an ongoing debate in the mental health field over whether mental disorder diagnoses should take contextual factors into account (Horwitz & Wakefield, 2007). In the *DSM-5* (APA, 2013), some context is required for the diagnosis of Posttraumatic Stress Disorder (PTSD); specifically, the experience of a traumatic event (e.g., killing a civilian in combat). Yet recent survey and interview data indicate that the onset of PTSD symptoms (e.g., exaggerated startle response) is not predicted most reliably by a traumatic event, but rather by a severe internal reaction (e.g., extreme fear; Boals & Schuettler, 2009). In the current work, we investigated the extent to which clinicians believe PTSD should be diagnosed in the context of a traumatic event, a severe reaction to an event, or both. In Experiment 1, we presented clinicians and clinical trainees with vignettes describing a hypothetical person’s PTSD behaviors preceded by a description of a traumatic or mildly stressful event and the person’s severe or mild reaction to the event. Participants then diagnosed each case. Given a traumatic event rather than a mildly stressful event, both clinicians and trainees made more diagnoses of PTSD, regardless of reaction type. This effect was replicated in a new sample of clinicians and clinical trainees Experiment 2. In Experiment 2, we also tested clinicians’ and trainees’ memory for events, reactions, and behaviors using a standard recognition task. Participants were less accurate for reactions than for events or behaviors, and tended to falsely recognize reactions that matched the events or behaviors in severity (and valence). Implications for clinical diagnostic systems and clinical conceptualizations of PTSD are discussed.

(266 words)
Events, Reactions, and Behaviors: Clinical Assessment of Posttraumatic Stress Disorder

Posttraumatic Stress Disorder (PTSD) is characterized by a stressful experience that leads to symptoms months later such as flashbacks, a heightened startle reflex, negative thoughts and feelings, impairment in daily functioning, and avoidance of stressor reminders (Diagnostic and Statistical Manual of Mental Disorders [DSM], 5th ed.; DSM-5; American Psychiatric Association [APA], 2013). An estimated 7.8% to 8.3% of adults in the U.S. develop PTSD in their lifetimes (Gabbay, Oatis, Silver, & Hirsch, 2004; Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; Kilpatrick et al., 2013). The central goal of the current work is to better understand how clinicians conceptualize, remember, and diagnose cases of PTSD, and to examine whether or not their reasoning aligns with that prescribed by the DSM-5 (APA, 2013).

The DSM and the special case of PTSD

The DSM is used in a number of countries (Mezzich, 2002); for example, clinicians in the U.S. must provide a DSM diagnosis for their clients to receive insurance reimbursement for mental disorder treatment (Kutchins & Kirk, 1997). This effectively constrains formal clinical diagnosis to the established framework of the DSM, which is designed to bring consensus amongst clinicians (DSM-5, APA, 2013; Kirk & Kutchins, 1992). DSM disorders are each defined by a set of rules: A list of diagnostic criteria and exclusions that specify whether a diagnosis should be made in any given case (DSM-5; APA, 2013). Because of ongoing controversies regarding the precise causes of disorders, the DSM has explicitly avoided specifying etiologies of disorders in its diagnostic criteria, beginning with the DSM-III (APA, 1980) and continuing through the text revision of the fourth edition of The Diagnostic and Statistical Manual of Mental Disorders (4th edition, text revision; DSM-IV-TR; American Psychiatric Association [APA], 2000). That is, for 33 years (1980-2013), the manual implied that
clinicians were expected to diagnose abnormal behavioral symptoms without being required to take into account the context in which those behavioral symptoms occurred. The current *DSM-5*’s (APA, 2013) introductory text informally acknowledges clinicians’ use of context to assess behaviors in diagnosis, but the vast majority of disorders in the *DSM-5* still do not explicitly include context in the diagnostic criteria (APA, 2013).

The diagnosis of PTSD, an exception to this rule, is a special case in the *DSM*. Although the diagnostic criteria for most *DSM* disorders have not included information about the context of behavioral symptoms for many years, PTSD is unique: The *DSM-III* (APA, 1980) up through the current edition (*DSM-5*; APA, 2013) have explicitly required context for the behavioral symptoms as one of PTSD’s diagnostic criteria. Further, whereas the behavioral symptoms required for a PTSD diagnosis have remained relatively stable across versions of the *DSM* containing PTSD, the definition of the stressor itself has undergone major changes in every version of the manual (Vasterling & Brewin, 2005).

In the most recent version of the *DSM*, the *DSM-5* (APA, 2013), the APA made a major modification to the PTSD diagnostic criteria regarding the stressor. The prior, fourth edition of The *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*; American Psychiatric Association [APA], 1994) and the subsequent text revision version (*DSM–IV–TR*; APA, 2000) had asked clinicians to diagnose PTSD only if the following conditions were met: The person was exposed to actual or threatened death, serious injury, or sexual violence (the A1 criterion), the person’s exposure to a traumatic event was accompanied by a reaction of intense fear, helplessness, or horror (the A2 criterion), and behavioral symptoms are present (e.g., intrusive, unwanted, and distressing memories of an event, avoidance of reminders of an event, exaggerated startle response, etc.; criteria B through H). In contrast, the current edition (*DSM-5*,
APA, 2013) removed the A2 criterion from PTSD diagnosis, such that clinicians are no longer being asked to attend to a formerly relevant part of PTSD etiology (APA, 2013). That is, at present, it includes only the following criteria for a PTSD diagnosis: the A1 criterion and criteria B through H.

The APA’s decision to remove the A2 criterion from PTSD diagnosis marked a culmination of a contentious and intense debate in the psychiatric literature on whether the A2 criterion legitimately should be used for PTSD diagnosis (Bedard-Gillgan & Zoellner, 2008; Hathaway, Boals, & Banks, 2010). Evidence was amassed that pointed in two fundamentally incompatible directions at once: Much work supported the A2 criterion’s inclusion; yet much work simultaneously refuted the A2 criterion’s utility.

Supporting the inclusion of the A2 criterion was survey and interview research demonstrating that an intense immediate reaction of fear, helplessness, or horror (A2), and not the experience of an objectively traumatic event (A1), predicts the onset of PTSD behavioral symptoms (Boals & Schuettler, 2009; Bodkin, Pope, Detke, & Hudson 2007; Gold, Marx, Soler-Baillo, & Sloan 2005; Rubin & Feeling, 2013). That is, regardless of whether a clinician perceives an event in a client’s life to be traumatic, it is the client’s intense immediate reaction to any given event that best predicts the onset of PTSD behavioral symptoms. Indeed, the vast majority of people who develop PTSD experience an intense immediate reaction at the time of a stressful event (e.g., 89% in Brewin, Andrews, & Rose, 2000). Further, several studies have shown that the more negative a person’s immediate reaction to a stressful event, the greater likelihood that person will develop subsequent disordered behaviors, a relationship which is predicted by individual cognitive characteristics (e.g. anxiety sensitivity, number of past traumas; Ehring, Ehlers, & Glucksman, 2006; Elwood, Hahn, Olatunji, & Williams, 2009; Riskind &
Alloy, 2006; Rosen & Lilienfeld, 2008). During the writing and development of the DSM-5 (APA, 2013), many papers were published that proposed revisions to PTSD diagnostic criteria to better align it with the above research (e.g., Spitzer, First, & Wakefield, 2007). Ultimately, these proposals were unsuccessful in shaping the criteria for PTSD in the DSM-5, as the APA decided to give greater weight to other work pointing in the opposite direction.

Refuting the utility of the A2 criterion in PTSD diagnosis, and cited by the APA as most important to their decision to remove the A2 criterion, was research demonstrating that the A2 criterion, although present in many clients who suffer from PTSD, does not increase the likelihood that someone will develop PTSD symptoms (i.e., diagnostic criteria B-H; Bedard-Gillgan & Zoellner, 2008). An interview study conducted on over 2000 people about their lifetime histories of traumatic events and subsequent PTSD suggested that there is a strong positive correlation between the presence of A1 and A2 (Breslau & Kessler, 2001), such that a small number of additional people would qualify for a PTSD diagnosis with the removal of A2 (Karam et al., 2010). Further, many people report non-A2 immediate reactions at the time of a traumatic event (e.g., anger, shame, numbness, or a military-trained response; Adler, Wright, Bliese, Eckford, & Hoge, 2008; Hathaway et al., 2010), and as such were not eligible to receive a DSM-IV-TR (2000) PTSD diagnosis, even if all other symptoms were met. In a 2013 pre-DSM-5 press release on changes made to PTSD diagnostic criteria, the APA stated that A2’s lack of independent predictive validity was the reason for not including the criterion in the DSM-5 (APA, 2013).

Thus, the exact kind of context needed for a PTSD diagnosis was altered from the DSM-IV-TR (APA, 2000) to the DSM-5 (APA, 2013), and not without contention. Given that clients’ intense reactions have shown a diverse set of influences on clinical researchers’ academic
assessments of PTSD diagnosis, it was an open and uncertain question as to what exact kinds of context influence clinicians’ diagnostic decisions for potential PTSD behaviors.

**Causal life-event context influences perceptions of mental disorder symptoms: A proportionate-response effect in clinical judgment**

Research on clinical judgment illuminates how clinicians perceive and conceptualize mental disorders, and how these views influence the diagnoses they make. For example, clinicians do not appear to represent mental disorders as a list of symptoms, the way mental disorders are presented in the *DSM*. Instead, clinicians represent mental disorders as symptoms that tend to participate in particular cause and effect relationships with one another that stem from a central cause (Kim & Ahn, 2002). This is a theory-based representation of mental disorders, wherein symptoms vary in importance for diagnosis based on their degree of centrality in the theory. Descriptions of hypothetical patients with symptoms that are causally central to a clinician’s concept of a disorder are more often diagnosed with a mental disorder and are considered better representations of the mental disorder than are descriptions of hypothetical patients with symptoms peripheral to the clinician’s concept of that mental disorder (Kim & Ahn, 2002). Further, clinicians are more likely to remember the most causally central symptoms than peripheral symptoms (Kim & Ahn, 2002). Additional research pointed to the special importance of the root cause in the representation of a disorder (Ahn, Marsh, Flanagan, & Sanislow, 2006). Clinicians who feel the root cause of a mental disorder is biologically based are more likely to recommend medication as treatment, whereas clinicians who feel the root cause of a mental disorder is psychologically based are more likely to recommend therapy as treatment (Ahn et al., 2009). In sum, the *DSM* does not represent mental disorders as symptoms of varying importance arranged in a theory containing one or more root causes, but clinicians do.
Because clinicians represent mental disorders as causal theories, they tend to use causal context to understand and diagnose mental disorders, in direct opposition to the DSM’s track record of (in most cases) not including causal context in disorder diagnostic criteria. Previous work has demonstrated that, when provided with causal context, clinicians do use it to understand disordered behaviors. In the domain of disorders, it has been shown that a plausible, causal, externally-controlled life-event explanation (e.g., the loss of a loved one; a severe car accident) for DSM-IV-TR symptoms of mental disorders (e.g., feelings of sadness and difficulty sleeping; overwhelming anxiety and difficulty concentrating; Ahn et al., 2003; Kim & LoSavio, 2009) gives clinicians the intuitive sense of understanding them (Kim et al., 2012a). This perception of understanding triggers the inference that the bizarre behaviors are less abnormal and less in need of treatment, even though those behaviors continue to cause the same significant levels of distress and disability as they would if not explained (Kim et al., 2012a). This growing body of evidence suggests that there is an understanding-normality effect in assessment, wherein a causal explanation tends to make an unusual event or behavior seem more prevalent, normal, and, in the case of disordered behaviors, even more psychologically healthy and less in need of treatment than disordered behaviors alone (Ahn, Novick, & Kim, 2003; Kahneman & Tversky, 1982; Kim & LoSavio, 2009; Kim, Paulus, Gonzalez, & Khalife, 2012a; Kim, Paulus, Nguyen, & Gonzalez, 2012b; Koehler, 1991; Meehl, 1973; Weine, Kim, & Lincoln, 2015).

Kim et al. (2012a) found support for a proportionate-response effect in clinical judgment. Generally, proportionality is perceived as one cue or indicator of a causal relationship between two or more elements or occurrences (Einhorn & Hogarth, 1986). For example, people may feel that a large red stain on a carpet was more likely caused by a large spilled glass of red wine than by a tiny spilled glass of red wine. In this case, the proportionality between the size of the stain
and the size of the glass leads to the judgment that the larger glass of wine caused the stain. In work by Kim et al. (2012a), clinicians judged disordered behaviors (i.e., DSM-IV-TR symptoms of PTSD or depression) that were placed in the context of a traumatic life event (e.g., experiencing a school shooting) or a mildly distressing life event (e.g., reading a news article about a school shooting across the country). Clinicians also read a description of someone’s disordered behaviors placed in one of the two contexts above. Clinicians found the cases where a traumatic life event led to disordered behaviors to be easier to understand and less psychologically abnormal than cases where a mildly distressing life event led to the exact same disordered behaviors. Similarly, they found cases where a mildly distressing life event led to unaffected behaviors to be easier to understand and less psychologically abnormal than cases where a traumatic life event led to the exact same unaffected behaviors. In sum, clinicians found cases where an event and behaviors were proportionate in strength and in valence to be easier to understand and less psychologically abnormal than cases where they were disproportionate (Kim et al., 2012a).

In the current work, we attempted to systematically examine seven questions across two experiments; Questions 1 and 2 are the central issues for this work, and Questions 3-7 supplement or follow from the first two. Questions 6 and 7 are investigated in Experiment 2, following the results of Experiment 1. First, all seven questions are introduced below.

**Question 1:** Are clinicians’ diagnoses of PTSD systematically influenced by the exact kinds of causal context required for a PTSD diagnosis in the DSM-5 (APA, 2013) or the DSM-IV-TR (APA, 2000)?

By “exact kinds” of causal context, we mean events versus reactions: In the remainder of this paper, the term *events* will be used to refer to an individual’s life experiences prior to the
onset of symptoms (i.e., the A1 criterion; *DSM-IV-TR*, APA, 2000; *DSM-5*, APA, 2013). The term *reactions* will be used to refer to an individual’s thoughts, feelings, and emotions occurring immediately in response to the event (i.e., the A2 criterion; *DSM-IV-TR*, APA, 2000).

This research question merits a full investigation independent of the previously discussed work documenting the effects of causal context on clinical judgments because it is unclear whether both types of causal context, events and reactions, influence clinical judgments about PTSD. In past work, it was found that clinicians take events into account in diagnosis judgments (Kim & LoSavio, 2009; Kim et al., 2012a; Kim et al., 2012b); however, whether and how clinicians make use of information about reactions has not yet been explored. Thus, Experiments 1 and 2 examined how clinicians use event and reaction context, separately and interactively, to understand, assess, and diagnose behavioral symptoms.

One reasonable hypothesis is that traumatic event context, but not intense reaction context, will influence clinical judgments of PTSD behaviors. Clinicians may simplify judgments by following a heuristic in which behaviors are understood best in the context of an initial causal link (the event) rather than a middle causal link (the reaction). Past research might suggest that people would be unlikely to discard the initial causal link, as they may view it as the root cause (the event; Ahn et al., 2009). Thus, people would be influenced only by plausible causal connections between events and behaviors when assessing other people’s behaviors. In addition, clinicians may deliberately align their diagnostic decisions with the most up-to-date version of the PTSD diagnostic criteria (i.e. in the modified *DSM-5*; APA, 2013), in which a traumatic event and behavioral symptoms (but not an intense reaction) must be identified for a PTSD diagnosis. Further, clinicians may make judgments in alignment with the APA’s decision that only the A1 criterion and not the A2 criterion should not be used in PTSD diagnosis, which
was based on work that has shown low predictive validity of the A2 criterion for the onset of
PTSD behavioral symptoms (e.g., Bedard-Gillgan & Zoellner, 2008). Any of the above could
reasonably be expected to lead clinicians to be solely influenced by the causal context of
traumatic events when reasoning about PTSD behavioral symptoms.

A plausible alternative hypothesis is that intense reaction context, but not traumatic event
context, will influence clinicians’ perceptions of PTSD behaviors. One possible reason for this is
that clinicians may find intense immediate psychological (emotional) reactions to be the most
salient kind of causal context when assessing disordered or distressed behavioral symptoms.
They may also pay closer attention to the reaction context than the event context because the
reactions most immediately temporally precede the symptoms and could be perceived as most
clearly responsible for bringing them about (Einhorn & Hogarth, 1986). Further, clinicians may
make judgments in alignment with the body of work demonstrating a strong link between intense
reactions (A2) and the onset of PTSD symptoms (e.g. Boals & Schuettler, 2009). Thus, clinicians
may only be influenced by the causal context of intense reactions when reasoning about PTSD.

An additional alternative hypothesis is that only the presence of both a traumatic event
and an intense reaction will influence perceptions of PTSD behaviors. One possible reason for
this is that clinicians might perceive the event, reaction, and behavior information as components
in a causally linked chain. As discussed earlier, past work by Einhorn & Hogarth (1986) suggests
that people may only form a causal chain between events, reactions, and symptoms when all are
proportionate in strength and negativity. If so, then people might assess PTSD behaviors only in
the context of a proportionately strong and negative traumatic event and intense reaction.
Further, clinicians may make judgments in alignment with the DSM-IV (APA, 1994) and DSM-
IV-TR (APA, 2000), in which both types of context have been needed for a PTSD diagnosis for
nearly 20 years. Again, the removal of intense reactions (A2) occurred only recently (in the DSM-5; APA, 2013). Therefore, practicing clinicians may rely on the established model of PTSD with which they have been making diagnoses for some time, to understand and diagnose PTSD behaviors.

**Question 2: Does this causal context increase or decrease clinicians’ tendency to make diagnoses of PTSD?**

We also examined whether such causal context *increases or decreases* clinicians’ tendency to give diagnoses of PTSD. One hypothesis is that clinicians will make more PTSD diagnoses for PTSD behaviors given causal context. Clinicians may be strongly influenced by the *DSM-IV-TR* (APA, 2000) and/or *DSM-5* (APA, 2013) frameworks, wherein causal context is required for a PTSD diagnosis. Thus, given this fact, it would also make sense to find that clinicians make more PTSD diagnoses for PTSD behaviors in causal context.

A plausible alternative hypothesis is that clinicians will make fewer PTSD diagnoses for PTSD behaviors given causal context. One reason for this is past work by Kim et al. (2012b), which revealed that clinicians judged Major Depressive Disorder (MDD) behaviors as less indicative of a MDD diagnosis in the context of a causal life-event, compared to the context of a neutral life-event. If this is a general tendency across disorders, we would expect clinicians to make fewer PTSD diagnosis responses for PTSD behaviors in causal context.

**Question 3: Do the above findings for PTSD extend to clinicians’ diagnostic judgments about disordered behavioral symptoms in general? Do they extend to clinicians’ diagnostic judgments about behaviors in general, including non-disordered ones?**

Our third research question set, also examined in Experiments 1 and 2, is concerned with whether any effects of causal context on clinicians’ diagnostic judgments for PTSD also extend
to their reasoning about other disorders in the DSM, reflecting a disorder-nonspecific effect. To conduct an initial examination of this issue, we asked participants to reason about MDD behaviors as a comparison disorder to PTSD because the *DSM-5* does *not* include event or reaction context in the MDD diagnostic criteria (APA, 2013).

One hypothesis is that causal context will affect clinicians’ diagnostic judgments of PTSD and MDD in the same way, indicating a possible disorder-nonspecific effect of causal context. Causal context may help clinicians explain away the need for a diagnosis for *any* disordered behaviors because it helps the disordered behaviors seem less difficult to understand. Or causal context, in the form of a severely negative stressor, may make it seem more likely that disordered behaviors would develop, leading clinicians to search more carefully for possible diagnoses matching those behaviors.

An alternative is that causal context affects clinicians’ judgments of PTSD and MDD differently, indicating that causal context does not affect judgments of all disorders in the same manner. For example, causal context may increase PTSD diagnoses but not affect MDD diagnoses, in line with the *DSM-5* (APA, 2013) diagnostic criteria for these disorders. Again, according to the *DSM-5* (APA, 2013) no context need be identified for a MDD diagnosis. If this is the case, future disorder-specific investigations should be conducted to examine how clinicians reason about the full spectrum of disorders.

We were also interested in whether causal context affects clinicians’ diagnostic judgments about non-disordered behaviors in the same way that it affects clinicians’ diagnostic judgments about PTSD, indicating a possible behavior-general effect. To examine this, we asked clinicians to reason about cases of mildly negative behaviors. Mild, non-disordered behaviors, which we henceforth refer to as distressed behaviors, appropriately contrast with disordered
behavioral symptoms because they describe someone coping with a stressful life event in a way that does not indicate a disorder (as opposed to someone exhibiting no reaction at all after a stressful life event; Schwartz & Link, 1989).

One hypothesis is that causal context will not affect clinicians’ diagnostic judgments of more ordinary behaviors in the same way. Indeed, previous work suggested that clinicians are not differently influenced by traumatic life event context and mildly stressful life event context when reasoning about distressed behaviors, as they judged distressed behaviors to be equivalently psychologically abnormal and easy to understand in both contexts (Kim et al., 2012a). In contrast, in that work, clinicians were differently influenced by traumatic and mildly stressful life event context for completely unaffected behaviors without any negative emotions at all, judging them to be more psychologically abnormal in traumatic life event context than in mildly stressful life event context.

An alternative is that causal context will have a similar effect on clinicians’ diagnostic judgments about a wider range of behaviors, including distressed behaviors. For example, if the event is sufficiently traumatic or the reaction sufficiently severe, clinicians may feel that the distressed behavioral symptoms are a sub-syndromal response that merits proactive treatment and therefore a diagnosis even before disordered symptoms develop. Given past findings (Kim et al., 2012a), it is unlikely that causal context would make all behaviors seem less in need of a diagnosis (e.g., by helping even distressed behaviors seem less difficult to understand).

**Question 4:** Is the proportionate-response effect found in past work for psychological abnormality judgments and difficulty-of-understanding judgments replicated, whether or not the effect extends to open-ended diagnosis judgments?
Again, Kim et al. (2012a) found clinicians’ judgments of psychological abnormality and difficulty understanding to align, such that proportionately negative traumatic event context made PTSD and MDD behavioral symptoms seem less difficult to understand and less psychologically abnormal. We expected to replicate this proportionate-response effect for PTSD and MDD behaviors in this study. We based this prediction on past work, whether or not these judgments align with open-ended diagnosis judgments in this study (for open-ended diagnosis hypotheses, please see Questions 1 and 2). Interestingly, if the proportionate-response effect is replicated for psychological abnormality and difficulty-of-understanding judgments, yet is not found in open-ended diagnosis judgments, this would suggest that diagnoses may not be made on the basis of the perceived psychological abnormality of behaviors (a possibility that is considered in much more detail in the General Discussion and Appendix E).

**Question 5: Do these judgments about PTSD, MDD and distressed behaviors differ across levels of clinical expertise?**

In Experiments 1 and 2, we explore whether there is any apparent effect of training on the degree to which causal context influences people’s diagnoses of behaviors. It is critical to understand how the diagnoses of clinicians—trainees and experts—are influenced by causal context. Formal clinical diagnoses are made by someone trained to do so; they help clinicians determine what types of treatment would be suitable, and in the U.S., allow clients to receive insurance reimbursement for treatment. Thus, in both experiments, we sampled both practicing clinicians and clinical trainees enrolled in a graduate program for clinical psychology.

One possibility is that the assessment decisions of clinical trainees and practicing clinicians are influenced by causal context in the same way, as in past work (Ahn et al., 2003), which demonstrated the understanding-normality effect in both clinical trainees and clinicians.
That is, both groups judged disordered behaviors as less abnormal when traumatic event context was provided than when it was not provided. However, this past work was conducted with artificial disorders and did not measure diagnosis judgments.

On the other hand, it is possible that expertise yields shifts in the influence of causal context. Between clinical trainees and practicing clinicians there are differences in relative amounts of clinical experience with clients, thinking through diagnoses, and using the different versions of the DSM (e.g., DSM-IV-TR, APA, 2000 versus DSM-5, APA, 2013). Trainees have begun to learn the DSM-5 (APA, 2013), just as clinicians have, but they have had comparatively less experience, if any, using the DSM-IV-TR (APA, 2000). Clinicians have spent time not only learning the DSM-IV-TR (APA, 2000), but also implementing it in their clinical practice and using it in diagnosis for years. In addition, in general, clinicians have accumulated not only much more experience and hours using versions of the DSM in clinical practice compared to trainees, but also thinking about and interacting with clients. For example, clinicians may be less likely to adhere to the newly updated DSM than trainees, as they have been applying the DSM-IV-TR (APA, 2000) in practice for up to 13 years, and may have used earlier versions from the DSM-III (APA, 1980) and on to diagnose PTSD using the A2 criterion for up to 33 years. We were able to explore any such differences in these experiments.

**Question 6: Are clinicians and clinical trainees more likely to have distorted memories of reactions than of events or behaviors?**

In Experiment 1, reaction context did not apparently influence any of the judgments we elicited, whereas event context and behavior type were clearly influential. A generally accepted indication of distorted remembering is when information not present is mistakenly remembered later as having been present (i.e., a “false alarm;” Deese, 1959; Roediger & McDermott, 1995).
In Experiment 2, we employed a standard recognition task, with event items, reaction items, and behaviors items from the vignettes in Experiment 1. For event, reaction, and behaviors items, we asked them to recognize original (correct) elements, false elements that were matched to the original ones in strength (e.g., mild versus severe) but not thematic content, and false elements that were matched to the original ones in terms of thematic content but not strength.

We hypothesized that clinicians and clinical trainees may be more likely to have distorted memories of reaction information than of event or behavior information when learning about a hypothetical person. Specifically, we predicted that clinicians and clinical trainees would make more false alarms for reactions than events or behaviors. This might occur if they perceive reaction information (either implicitly or explicitly) as less important than, or secondary to, event or behavior information. Events and behaviors, being more directly observable, might be more salient in real life and therefore more likely to be encoded precisely than reactions, which unfold internally (to the patient). Thus, even if reaction information is described clearly and presented saliently, participants may by default fail to encode it and instead infer it indirectly from their memories of the events and behaviors. Furthermore, clinicians and clinical trainees may have better memories of behaviors than reactions because behaviors are the basis of almost all diagnoses in the *DSM* (in all versions since the *DSM-III*; APA, 1980), whereas reactions were included as diagnostic criteria only in the rare case of PTSD in only the *DSM-IV* (APA, 1994) and *DSM-IV-TR* (APA, 2000).

An alternative possibility is that clinicians’ and clinical trainees’ memories for events, reactions, and behaviors do not differ. Accordingly, we might expect to find false alarm rates to be equivalent for events, reactions, and behaviors, implying that participants were not influenced
by reaction context when making diagnosis judgments for a different reason (e.g., because they found it to be irrelevant to the judgments, not because their memories of it were distorted).

**Question 7: Are clinicians and clinical trainees susceptible to forming distorted memories of reactions such that their strength of negative valence is proportionate to that of the other elements?**

Given past work on false memory formation, we expect that people will form some distorted memories in terms of content (Roediger and McDermott, 1995), but whether they make systematic distortions in terms of strength of negative valence is an open question. In disproportionate cases (i.e., cases in which the reaction information is disproportionate in the strength of negative valence to the event, behaviors, or both), we hypothesize that clinicians and clinical trainees might be particularly susceptible to forming distorted memories of reactions in a manner that would better reconcile them with the event or behaviors. That is, they may retrieve the event and behavior information from memory relatively directly, but more indirectly infer the reaction information under the assumption that it is most likely to have been proportionate to the event or behaviors in strength of negative valence. In proportionate cases, clinicians and clinical trainees may do the same, but because proportionate cases already match in strength of negative valence, the inferred reaction memory will be correct. Overall, such findings would imply that participants did not appear to use reaction information in Experiment 1 at least in part because they were especially susceptible to forming distorted memories of it. An alternative possibility is that clinicians and trainees will not distort the reaction information to better align with the event or behaviors in terms of strength of negative valence; that is, they may distort it, but not in any systematic way.
Experiment 1: Contextual Influences on Open-Ended Diagnoses of PTSD Behaviors

To examine Questions 1-5, we asked licensed, practicing clinical psychologists and clinical graduate students to assess hypothetical individuals described in short vignettes. The vignettes each described PTSD, MDD, or subclinical distressed behavioral symptoms in the context of events and reactions to those events. Participants were asked to make diagnoses for each vignette, to judge the psychological abnormality of the person, and to judge their own difficulty understanding that person’s behaviors.

Methods

The Northeastern University Internal Review Board approved all experiment protocols.

Participants.

Clinicians. Fifty-six licensed practicing clinicians with Ph.D. degrees (36 female) participated in response to a mailed postcard invitation to an online study. A random sample of 240 clinicians from 47 US states was taken, via psychologytoday.com (a publicly available online directory of U.S. mental health professionals), to receive postcards. In an attempt to increase the diversity of our sample, an additional 38 clinicians listed in the Directory of Ethnic Minority Professionals in Psychology (4th ed.; APA, 2001) were also sampled. Participating clinicians were offered a $25 gift card to an online retailer. A follow-up post card was sent two to four weeks later to clinicians who did not respond to the first mailing. We separately verified that each clinician had a license in good standing in his or her state of practice. Because Alabama, Arkansas, and Minnesota did not offer free clinician licensure information online, clinicians from these three states were not sampled. Out of the 278 clinicians who were mailed postcards, 11 clinicians’ cards were returned because the clinician had moved, yielding a 21.0%
response rate, which is comparable to other work using postcard recruitment to sample a clinician population (Rottman, Ahn, Sanislow, & Kim, 2009).

Clinicians reported a mean age of 50 years (range: 26-71) and most self-identified as White (95%), with the remaining participants identifying themselves as Asian (3%) and Black (2%). In addition, 7% self-identified as Hispanic. Eighty-nine percent held a private practice. Clinicians had a median of 19 years of clinical experience (interquartile range: 12 years, 28 years), and held a range of primary clinical orientations (47% Cognitive or Cognitive-Behavioral, 18% Psychoanalytic, 18% Eclectic, 5% Humanistic, 5% Family Systems, 2% Behavioral, and 5% Other).

**Graduate students.** Eighty graduate students (60 female) were also recruited. Graduate students reported a mean age of 27 years (range: 23 - 35), and self-identified as White (88%), Asian (5%), Black (1%), and multiple races (4%; the remaining participants declined to report race). In addition, 8% self-identified as Hispanic. Graduate students were enrolled in APA-accredited clinical psychology graduate programs in the U.S., and were contacted by email using addresses publicly available on their department websites. For participation, graduate students were offered entry to a raffle for one of several $25 gift cards to an online retailer.

Graduate students in the participating sample provided psychotherapy services in a variety of settings (39% psychology department clinics, 18% hospitals, 16% other, 11% community clinics, and 16% no service experience), and had a median of 3 years of clinical experience (interquartile range: 0.88 years, 4 years). Graduate students held a range of primary clinical orientations (69% Cognitive or Cognitive-Behavioral, 14% Behavioral, 6% Eclectic, 5% Psychoanalytic, 6% Other).
**Materials.** We adapted five sets of vignettes from Kim et al. (2012a) and created one, for a total of six sets of vignettes that each described a hypothetical person. There were three parts of each vignette set: The event, the reaction, and the behaviors. A paragraph labeled “Recent Experiences” included a description of the event and a description of the reaction, roughly equated to each other for length. The event and reaction were both included in the same paragraph to meet the *DSM-IV-TR* (APA, 2000) A2 criterion (e.g., to specify that the A2 reaction was in response to the A1 event; please see Table B1 for a partial sample vignette, and Appendix A for full vignettes). A paragraph labeled “Subsequent Behaviors” included five behavioral symptoms that described PTSD (taken from Criteria B-H), Major Depressive Disorder (MDD), or distressed behaviors.

The topics of the events described in the vignettes were either modified from Kim et al. (2012a), drawn from the *DSM-IV-TR* casebook (Spitzer, Gibbon, Skodol, Williams, & First, 1994), or drawn from published studies of people’s experiences leading to PTSD (Matsakis, 1996; Rothschild, 2000). Each event was three to five sentences long, and included the following: A brief description of the person in the vignette (e.g. “Lucy teaches at an elementary school…”), a detailed description of a particular occurrence (e.g. “Lucy and her students were pinned down by sniper fire.”), the person’s whereabouts at the time (e.g. “…on the playground”), experiences during the event (e.g. “Lucy saw a child get shot.”), and / or actions at the time (e.g. “She tried to hide.”). Each event also included a brief description of a resolution, as cue that the event was a single finite occurrence and not a continuing experience (e.g. “After the gunfire ceased…”). See again Table B1 for a partial sample vignette and Appendix A for full vignettes.

The events were either traumatic or non-traumatic (everyday events). *Traumatic events* were defined by strict adherence to the A1 criterion from both the *DSM-IV-TR* and *DSM-5*, and
described an objectively traumatic event involving actual or threatened death, serious injury, or sexual violence, which was experienced by a person directly or indirectly, or happened to a person’s family member or friend (e.g. receiving an injury in a severe car accident; *DSM-IV-TR*, APA, 2000; *DSM-5*, APA, 2013). *Everyday events* were designed to fall markedly short of meeting either the *DSM-IV-TR* (APA, 2000) or *DSM-5* (APA, 2013) PTSD A1 criterion for a traumatic event (e.g., did not describe an objectively traumatic event). Four of the everyday events described a less severe form of its corresponding traumatic event vignette (e.g. experiencing a fender-bender); two of the everyday events described a secondhand experience of its corresponding traumatic event vignette (e.g. reading about a bad car accident in the newspaper).

The reactions were described as being either intense or mild. The reaction text was generated specifically for this study, and tailored to each event, such that each of the two reaction vignettes were interwoven coherently with each of the two event vignettes (e.g. “They were pinned down by sniper fire” [event information], “Lucy felt paralyzing fear” [reaction information], “She hid behind a trashcan” [event information], “Her thoughts raced” [reaction information]). The reactions were described with roughly the same number of words as events, and included the person’s thoughts (e.g., “she realized that safety was an illusion”), feelings (e.g., “she felt like she had failed her students”), and emotions (e.g., “she experienced an extremely heightened sense of fear”) that occurred within an individual at the time of the event. Reaction information was also carefully written to avoid inadvertently describing a behavioral symptom of one of the disorders. For example, the reaction could not be described as someone feeling jumpy, because this could potentially be interpreted as indicating the presence of the PTSD criterion E, a heightened startle reflex (*DSM-5*, APA, 2013).
Intense reactions were defined by strict adherence to the A2 criterion in DSM-IV-TR (APA, 2000), as a severe internal reaction of fear, helplessness, or horror that encompassed someone’s thoughts, feelings, and emotions at the time of an event (e.g. feeling afraid and thinking about fleeing; DSM-IV-TR, APA, 2000). Mild reactions did not meet the DSM-IV-TR A2 criterion (e.g., were not an intense reaction of fear), and described someone’s mildly negative or neutral thoughts and feelings at the time of an event (e.g., feelings of surprise and thoughts of concern; APA, 2000). Please see Table B1 for a partial sample vignette and Appendix A for full vignettes.

The final paragraph in each vignette, labeled “Subsequent Behaviors,” appeared after the paragraph labeled “Recent Experiences.” The “Subsequent Behaviors” for each vignette comprised one of three sets of behavioral symptoms. One set described PTSD (and no other disorder, including MDD). A second set described Major Depressive Disorder (MDD, and no other disorder, including PTSD). A third set of “subsequent behaviors” described mild, non-disordered behaviors, as a control to both sets of disordered behavioral symptoms.

Each set of disordered “Subsequent Behaviors” included five symptoms that satisfied DSM-IV-TR and DSM-5 diagnostic criteria for either PTSD or MDD. Each set of distressed “Subsequent Behaviors” included five mild, non-disordered behaviors that did not meet diagnostic criteria for any disorder. Each behavioral vignette was between 100 and 120 words; length was roughly equated between the different versions for each vignette. Each behavioral vignette described behaviors that continued for more than one month after the “Recent Past” to meet DSM-IV-TR (APA, 2000) and DSM-5 (APA, 2013) criteria that PTSD symptoms continue for more than one month after the event, and to differentiate behavioral vignettes from
Adjustment Disorder$^2$ and Acute Stress Disorder, in which symptoms begin and remit within one month after an event (and in the case of Acute Stress Disorder, last no more than one month; APA, 2013).

**Manipulation Check.** Sixteen undergraduate psychology students (10 female) independently rated each event for how traumatic it seemed on a Likert scale of 1-9, where 1=not at all traumatic and 9=extremely traumatic, and each reaction for how severe it seemed on a Likert scale of 1-9, where 1=not at all severe and 9=extremely severe, in randomized order. Because these manipulation check participants were rating two events and two reactions for each vignette, which were labeled with only one hypothetical person’s name (e.g., Lucy), we needed to ensure that they were assessing each vignette individually. Thus, for the purposes of the manipulation check only, vignette names were substituted with other names matched for age, gender, and country of origin (Social Security Administration, 2010; e.g., Lucy’s two events in Table B1 were presented with the names Lucy and Megan; Lucy’s two reactions were presented with the names Stephanie and Heather). Paired samples t-tests revealed that participants found traumatic events ($M = 7.52, SD = .77$) to be perceived as more traumatic than everyday events ($M = 3.67, SD = 1.34; t(15) = 9.15, p < .001$), and intense reactions ($M = 7.17, SD = 1.04$) to be perceived as more severe than mild reactions ($M = 3.55, SD = .96; t(15) = 12.14, p < .001$). Further, all means ran in this same direction for each vignette.

---

$^2$ Adjustment Disorder is characterized by anxiety, poor ability to concentrate, negative mood (e.g., sadness), worrying, and insomnia, among other symptoms. Notably, it develops in response to an event or situation, with symptoms appearing within zero to three months of exposure (DSM-IV-TR, APA, 2000; DSM-5, APA, 2013). Conversely, PTSD develops within one to six months of exposure.
**Design and procedure.** Each participant read about one of the two event types (traumatic or mild) combined with one of the two reaction types (intense or mild) for each vignette. Event type and reaction type were manipulated between-subjects. Behavior type was manipulated within-subjects; each participant read two vignettes describing behavioral symptoms of PTSD, two vignettes describing symptoms of MDD, and two vignettes describing distressed behaviors for a total of six vignettes.

There were three dependent variables: open-ended diagnoses, judgments of psychological abnormality, and judgments of difficulty understanding the “Subsequent Behaviors” described in each vignette. For the open-ended diagnosis question for each vignette, participants were asked to type in up to three diagnoses, and were explicitly instructed that one was required, which could be “none.” They were instructed to rate the likelihood of each diagnosis on a scale of 1-9, where 1 = very unlikely and 9 = very likely (please see Table B2 for exact text). They were asked to make judgments of psychological abnormality on a scale of 1-9, where 1 = very psychologically normal and 9 = very psychologically abnormal, and judgments of difficulty understanding the subsequent behaviors on a scale of 1-9, where 1 = very easy to understand and 9 = very difficult to understand (please see Table B2). All participants also answered demographics questions, including an open-ended question about their experience with the *DSM-5* (APA, 2013) to date (“What, if any, is your experience with the *DSM-5*?”). All clinical trainees completed the study in February 2014 and all clinicians completed the study between February and June 2014, nine to 13 months after the *DSM-5* (APA, 2013) was released.

After beginning the online survey, participants first read and signed a consent form, and then were given basic instructions. Participants read a total of six vignettes, and answered the same dependent variable for all six before moving on to answer another dependent variable for
all six, until all three dependent variables had been presented for each vignette. The order of dependent variables was randomized, and the order of vignettes presented for dependent variable was also randomized. After rating the vignettes, participants answered demographics questions, including the DSM-5 experience question. Participants then provided an email address to receive an e-gift card as a token of thanks (optional), and were debriefed.

Results and Discussion

Again, our two primary goals were to investigate what exact kinds of context – events, reactions, or both in conjunction – affect clinical perceptions of PTSD (Question 1), and to uncover the direction in which judgments were influenced by such context (Question 2). To do so, we examined whether event context, reaction context, expertise (i.e., clinicians vs. clinical trainees), and self-reported DSM-5 experience affects PTSD diagnoses for PTSD behaviors (the analyses of primary interest). Then, we also examined whether and how the same factors affect depression diagnoses for MDD behaviors and indications that no diagnosis was needed for distressed behaviors (Question 3). Please see Appendix D for secondary analyses on the open-ended diagnosis judgments not pertaining directly to our central research questions. In each case, the critical comparisons were between the different conditions for the same behaviors (e.g., traumatic event vs. everyday event). We explored whether, and how, event context and reaction context affects judgments of psychological abnormality and difficulty understanding for PTSD behaviors, MDD behaviors, and distressed behaviors (Question 4). Finally, throughout all of the above analyses, we checked for effects and interactions involving expertise (Question 5).

DSM-5 experience coding. We first coded for participants’ reported DSM-5 exposure, to enable us to determine whether it had any effect on diagnoses. Participant responses to the open-ended question (“What, if any, is your experience with the DSM-5?”) were coded by two
independent coders into three categories: no experience (e.g., “haven’t looked at it”), minimal experience (e.g., “a little”), and some experience (e.g., “a couple of workshops”), with 100% agreement between coders. Overall, 58.93% of clinicians and 45.0% of clinical trainees reported either no or minimal experience with the DSM-5 (APA, 2013) at the time of the study. Please see Figure B1.

**Open-ended diagnosis responses.** A full discussion of the range and frequency of all the various open-ended diagnosis responses made by the participants can be found in the “Other disorders identified” section below and in Appendix D. We ran twelve chi-square analyses to examine effects of event context, reaction context, expertise, and DSM-5 experience on diagnoses. Three of these analyses examined the effect of event context, three the effect of reaction context, three the effect of expertise, and three the effect of DSM-5 experience on the following diagnoses: PTSD diagnoses made for PTSD behaviors, depression or MDD diagnoses made for MDD behaviors, and indications that no diagnosis was necessary for distressed behaviors. For each behavior type, the relevant diagnosis was included in the analysis only if it was the only diagnosis made for the vignette, if it was rated as the most likely diagnosis for the vignette, or if it was tied for most likely diagnosis for the vignette.

**PTSD diagnosis responses.**

**PTSD diagnosis responses for PTSD behaviors.** Our primary questions were (Q1) whether clinicians were influenced by the causal context of events, reactions, or both in making diagnosis responses, and (Q2) whether this context increased or decreased the likelihood of PTSD diagnoses. Simultaneously, we also examined whether judgments about PTSD differed across levels of clinical expertise (Q5). Participants made more PTSD diagnoses for PTSD behaviors in the context of a traumatic event (75.74%) than in the context of an everyday event
(47.06%; χ²[2, N = 136] = 21.43, p < .001). Please see Figure B2. However, reaction context did not affect the number of PTSD diagnoses for PTSD behaviors (χ²[2, N = 136] = .79, p = .704); neither did expertise (χ²[2, N = 136] = 1.64, p = .440) or DSM-5 experience (χ²[4, N = 136] = 2.75, p = .601). In Experiment 2, we were able to replicate this finding in a separate sample of clinicians and clinical trainees (please see p. 115 for Experiment 2 analyses).

The PTSD diagnoses for PTSD behaviors made by participants were affected solely by event context, falling in line with the DSM-5 (APA, 2013) diagnostic criteria for PTSD, which indicate that an APA-defined traumatic event (i.e., a traumatic event that fulfills the A1 criterion) must be identified for someone to be diagnosed with PTSD. As participants’ PTSD diagnoses for PTSD behaviors were influenced by event context, we also wanted to know whether such context increased or decreased these diagnoses. We found that event context increased PTSD diagnoses. Thus, traumatic event context did not explain away the need for a PTSD diagnosis.

Participants’ PTSD diagnosis judgments were not affected by reaction context. Thus, their judgments did not align with the DSM-IV-TR’s (APA, 2000) diagnostic criteria, which indicate that both a traumatic event (A1) and an intense reaction (A2) must be identified for someone to be diagnosed with PTSD. Furthermore, participants did not diagnose PTSD after an intense reaction alone, even though compelling work (e.g., Boals & Schuettler, 2009) has suggested that intense reactions lead to PTSD symptomology, rather than events that others perceive to be traumatic.

There were no effects or interactions involving either expertise or the level of reported DSM-5 experience on participants’ PTSD diagnoses. Although clinicians and clinical trainees differed in number of years of clinical experience and distribution of primary theoretical orientation, these two groups did not differently diagnose PTSD for PTSD behaviors. Also, not
all clinicians and trainees were familiar with the new *DSM-5* (APA, 2013) criteria (please see Figure B1); however, their judgments generally aligned with the new edition of the manual regardless. As some participants knew the *DSM-5* (APA, 2013) criteria, they might have made diagnoses in line with it deliberately, but the rest could not have done so. It is possible that clinicians and clinical trainees intuitively think that PTSD occurs in the context of a traumatic event, rather than in the context of an intense reaction or both a traumatic event and intense reaction in conjunction, leading them to diagnose PTSD more often in the presence of an objectively traumatic life event than when such context is absent.

Given that all vignettes were presented in randomized order, we also needed to check whether the effects we found held true for only the first PTSD vignette that people diagnosed (i.e., when judgments were uncontaminated by any previous judgments). We ran four chi-square analyses to examine effects of event context, reaction context, expertise, and *DSM-5* experience on PTSD diagnoses made by participants who diagnosed a PTSD vignette first. Participants still diagnosed PTSD more frequently in traumatic event context than in everyday event context ($X^2 [2, N = 47] = 7.88$ $p = .019$), and were unaffected by reaction context, expertise, and *DSM-5* experience (all $ps \geq .338$) even when they first made open-ended diagnoses for a PTSD vignette.

*PTSD diagnosis responses for MDD and distressed behaviors.* Question 1 (whether clinicians were influenced by the causal context of events, reactions, or both in making diagnosis responses) and Question 2 (whether this context increased or decreased the likelihood of PTSD diagnoses) can be still further addressed by examining participants’ diagnoses of PTSD for vignettes describing MDD behaviors and distressed behaviors. These analyses can tell us what kind(s) of causal context, if any, elicited more PTSD diagnosis responses even for behaviors that were not characteristic of PTSD. Thus, we ran separate chi-square analyses to examine the
effects of event context, reaction context, expertise, and DSM-5 experience on PTSD diagnoses for MDD behaviors and PTSD diagnoses for distressed behaviors. Clinicians made more PTSD diagnoses for MDD behaviors in the context of a traumatic event (18.38%) than in the context of an everyday event (4.41%; $\chi^2[2, N = 136] = 10.73, p = .005$). Also, they made marginally more PTSD diagnoses for MDD behaviors in the context of a severe reaction (16.91%) than in the context of a mild reaction (5.88%; $\chi^2[2, N = 136] = 5.58, p = .061$), although these percentages of PTSD diagnoses for MDD behaviors in traumatic event context (18.38%) were considerably lower than the percentage of PTSD diagnoses obtained for PTSD behaviors in traumatic event context (75.74%). Expertise did not affect the number of PTSD diagnoses for MDD behaviors ($\chi^2[2, N = 136] = .79, p = .704$); nor did DSM-5 experience ($\chi^2[4, N = 136] = 2.75, p = .601$). Thus, both types of severely negative context (event and reaction) motivated clinicians and clinical trainees to view MDD behaviors as having PTSD, compared to cases without such severely negative context.

For distressed behaviors, participants also made marginally more PTSD diagnoses in the context of a traumatic event (4.41%) than in the context of an everyday event (0.73%; $\chi^2[1, N = 136] = 3.76, p = .052$). However, these percentages were very low in comparison to PTSD diagnoses for PTSD behaviors (and even very low compared to PTSD diagnoses for MDD behaviors). Also, reaction context did not affect the number of PTSD diagnoses for distressed behaviors ($\chi^2[1, N = 136] = .95, p = .698$); neither did expertise ($\chi^2[1, N = 136] = .01, p = .926$) nor DSM-5 experience ($\chi^2[2, N = 136] = .63, p = .731$). Thus, clinicians and clinical trainees alike made more PTSD diagnoses for merely distressed behaviors in traumatic event context than in everyday event context, although they did so very infrequently overall.
PTSD diagnosis responses: Summary. In summary, clinicians made more PTSD diagnoses for all three behavior types (PTSD, MDD, and distressed) when those behaviors were presented in the context of a traumatic event compared to the context of an everyday event. This demonstrates that traumatic event context is important to clinicians’ and clinical trainees’ concepts of PTSD, so much so that the presence of such context increases PTSD diagnosis responses even if the behavioral symptoms themselves do not fit the diagnostic criteria.

Depression diagnosis responses for MDD behaviors. Next, we examined Question 3, whether clinicians’ elevated rate of PTSD diagnoses in traumatic event context relative to everyday event context extends to their diagnostic judgments for MDD and non-disordered, distressed behaviors. First, we examined depression diagnoses for MDD behaviors. Clinicians’ diagnoses of depression for MDD behaviors did not differ by event context ($X^2[2, N = 136] = .90, p = .637$; please see Figure B2), by reaction context ($X^2[2, N = 136] = 4.38, p = .112$), by expertise ($X^2[2, N = 136] = 2.39, p = .302$), or by DSM-5 experience ($X^2[4, N = 136] = 2.57, p = .631$). Overall, participants diagnosed depression for 82.35% of MDD vignettes.

We replicated these findings in Experiment 2 with one exception. In Experiment 2, clinical trainees diagnosed depression for MDD behaviors more often than clinicians did. However, both groups still often diagnosed depression for MDD behaviors, and neither were influenced by event or reaction context (please see p.115 for analyses and discussion).

We also checked whether the results held true when examining only the diagnoses of people who saw the MDD vignette first. To do so, we ran four chi-square analyses to examine effects of event context, reaction context, expertise, and DSM-5 experience on MDD diagnoses made by participants who diagnosed a MDD vignette first. Participants were still unaffected by
event context, reaction context, expertise, and DSM-5 experience (all $p$s $\geq .440$) even when diagnosing MDD behaviors first.

Although clinicians’ and trainees’ depression diagnoses for MDD behaviors were not affected by event or reaction context, they nevertheless identified depression as an appropriate diagnosis for MDD behaviors the majority of the time. We had thought that these judgments would have aligned with past work, in which clinicians advocated a depression diagnosis less strongly for MDD behaviors in the context of a negative life-event than in the context of a neutral life event or no life event (Kim et al., 2012b). Yet, the current study differs from this past work in a number of ways. First, Kim et al. (2012b) deliberately did not ask participants to diagnose MDD, to avoid criterion counting; instead, the experimenters asked about depression more broadly, which could have included depression diagnoses not often elicited in this study (e.g., depressive disorder not otherwise specified). Second, the number of depression behaviors described in Kim et al. (2012b) were designed to fall just short of diagnostic criteria (i.e., the vignettes described fewer than five criteria and five are required for a MDD diagnosis), whereas the MDD vignettes in the current study described five MDD diagnostic criteria, meeting the threshold. Third, clinicians may be differently influenced by context when making open-ended diagnoses for MDD behaviors as opposed to simply rating whether a particular diagnosis should be made, because they take different information into account when formulating their own diagnoses rather than judging whether a particular given diagnosis is appropriate. Fourth, we might not have replicated this past work because the control conditions differed between this study and the past work. The mildly negative everyday event context in this experiment (e.g., witnessing a fender bender) may have been somewhat more negatively valenced than the neutral life events used in past work (e.g., taking one’s family to an activity at the local community
center; Kim et al., 2012b). If so, it is reasonable to expect any difference between the experimental and control conditions in the current experiment to be smaller than that found in this past work.

**No-diagnosis responses for distressed behaviors.** We further examined Question 3, whether the above findings for PTSD extended to clinicians’ judgments for a broader range of behaviors, by examining diagnoses for non-disordered behaviors. Participants made fewer indications that no diagnosis was necessary for distressed behaviors in the context of a traumatic event (79.41%) than in the context of an everyday event (90.44%; $X^2[2, N = 136] = 12.86, p = .002$). That is, the traumatic event made participants more likely, relative to an everyday event, to give some kind of disorder diagnosis. Please see Figure B2. Of their responses for distressed behaviors in traumatic event context other than “none” (20.59%), only 2.94% were PTSD diagnoses; the rest were other disorder diagnoses. Participants’ indications that no diagnosis was necessary for distressed behaviors did not differ by reaction context ($X^2[2, N = 136] = 1.07, p = .587$), expertise ($X^2[2, N = 136] = .406, p = .816$) or DSM-5 experience ($X^2[4, N = 136] = 4.29, p = .368$). Overall, participants indicated “none” for 84.93% of distressed behavior vignettes. That is, clinicians and trainees were able to identify that no disorder was present, when this was the case, most of the time.

In Experiment 2, we also found that participants indicated “none” for distressed behaviors often. However, unlike in Experiment 1, we did not find that they were influenced by context when doing so (please see p. 116 for results and discussion).

Interestingly, the above findings are consistent with the proportionate-response hypothesis (Kim et al., 2012a), wherein cases of life-events and behaviors that are proportionate in strength and valence (negativity or positivity; e.g., an everyday event and distressed
behaviors) were rated as less psychologically abnormal and less difficult to understand compared to life-events and behaviors are that are disproportionate in strength and valence (e.g., a traumatic event and distressed behaviors), which are rated as comparatively more psychologically abnormal and more difficult to understand. Notably, participants in this study diagnosed merely distressed behaviors with a disorder more often in disproportionate traumatic event context than in proportionate everyday event context. This finding suggests that the proportionate-response effect may occur in open-ended diagnosis judgments of distressed behaviors, in addition to the judgments studied in past work (i.e., abnormality and difficulty understanding judgments; Kim et al., 2012a).

Finally, to examine “none” responses for distressed behaviors that were uncontaminated by other diagnoses, we ran chi-square analyses to examine the effects of event context, reaction context, expertise, and DSM-5 experience on “none” diagnoses for distressed behaviors when participants diagnosed a distressed vignette first. Just as in the analyses including all “none” diagnoses for distressed behaviors, participants indicated “none” for distressed behaviors less frequently in traumatic event context (84.09%) than in everyday event context (93.48%, X²[2, N = 45, = 6.48, p = .039). Also, like in the main analyses, participants were unaffected by reaction context, expertise, and DSM-5 experience when making this judgment (all ps ≥ .224).

Other disorders identified. To give a broader picture of participants’ responses, not all disorder identifications constituted the appropriate response for each vignette with respect to the DSM-IV-TR (APA, 2000) or DSM-5 (APA, 2013; please see Table B3). Given space to provide up to three disorder responses, participants indicated a wide range of DSM disorders. In total, they suggested 24 different DSM disorders (please see Table B3 for response counts by behavior type, and Appendix D for a more complete description of the results of the open-ended disorder
identification task.) The majority of DSM diagnoses besides those of interest were adjustment disorder, which was diagnosed a total of 78 times across all 816 vignette presentations in this study (10%, please see Table B3), anxiety disorder (38 times, 5%), acute stress disorder (37 times, 5%), and generalized anxiety disorder (24 times, 3%). Please see Table B3 for other DSM diagnoses by behavior type. Clinicians and clinical trainees also diagnosed depression for 58 out of 272 PTSD vignettes (21%) and PTSD for 73 out of 272 MDD vignettes (27%). Participants indicated that no disorder was present for 9 out of 272 PTSD vignettes (3%), only 2 out of 272 MDD vignettes (<1%), but 233 out of 272 distressed vignettes (86%). Thus, the most common responses on our open-ended measure across the three behavior types were PTSD, depression, adjustment disorder, anxiety disorders, acute stress disorder, and indications that no disorder was present.

The diagnoses of PTSD, adjustment disorder, and acute stress disorder are likely a result of the recent past information described for each vignette. Indeed, a recent traumatic stressor is part of the DSM-IV-TR (APA, 2000) and DSM-5 (APA, 2013) criteria for PTSD, adjustment disorder, and acute stress disorder. At the very least, the recent past information that participants read described a recent stressor that might plausibly lead to distressed behaviors (i.e., the everyday event), so clinicians and clinical trainees may have judged even the everyday events to be traumatic enough to qualify someone for these diagnoses. However, the materials in this study were carefully constructed to avoid confusion between PTSD and adjustment disorder and acute stress disorder. Specifically, all behavioral symptoms (PTSD, MDD, and distressed behaviors) were described as persisting for more than three months past an event, or beginning more than three months after the event. In the DSM-IV-TR (APA, 2000) and DSM-5 (APA, 2013), adjustment disorder ought to be diagnosed only within three months of a single stressor and
acute stress disorder ought to be diagnosed only within four weeks of a single stressor. Therefore, technically neither disorder was described in any of the vignettes. However, it appears that clinicians and clinical trainees still felt that these two other diagnoses were possible for some vignettes, even though they did not satisfy the required time course exactly. Acute stress disorder shares many symptoms with PTSD (e.g., avoidance of trauma reminders, heightened startle reflex). Adjustment disorder, PTSD, and MDD share common symptoms of distress and impairment. Therefore, these alternative diagnoses are not implausible. Also, they were made far more infrequently than diagnoses of PTSD or MDD.

Anxiety disorder and generalized anxiety disorder were also diagnosed a smaller number of times (38 and 24 times, respectively). These disorders are characterized by anxiety symptoms that cause distress and impairment. Indeed, these two disorders were diagnosed most often for PTSD vignettes (44 times), which described disorder-appropriate anxiety symptoms, compared to MDD behaviors (9 times) or distressed behaviors (9 times). For PTSD behaviors specifically, anxiety disorder and generalized anxiety disorder were diagnosed more frequently in everyday event context (28 times) than in traumatic event context (16 times), suggesting that participants diagnosed these disorders more frequently when they recognized all of the symptoms for PTSD except for the traumatic event.

Only 2% of all responses were neither a DSM-IV-TR (APA, 2000) or DSM-5 (APA, 2013) disorder nor a “none” response (please see Table B3 for details by vignette and Appendix D for all responses). In total, clinicians and clinical trainees made 3 trauma-related responses (e.g., “exposure to disaster, war, or other hostilities”), 7 that resembled a DSM disorder or proposed DSM disorder (e.g., “adjustment-like disorder;” “complicated bereavement disorder”), and 3 states (e.g., “bereavement;” “grief”). It is not necessarily the case that these relatively rare
non-DSM diagnoses should be viewed as outside the realm of clinical utility. Most were approximations of an existing DSM disorder, or a proposed disorder that has not yet been included in the DSM. As such, these infrequent indications may help clinicians and trainees pinpoint effective treatment strategies in practice. Future work is needed investigate this possibility, preferably in a more ecologically valid setting.

**Psychological abnormality judgments and difficulty understanding judgments.** We then considered Question 4, whether the proportionate-response effect found in past work for abnormality judgments and difficulty understanding judgments was replicated in the current work. This question was particularly important to explore given that the diagnosis judgments for PTSD and MDD did not run in the predicted direction. For psychological abnormality judgments and difficulty understanding judgments, we ran three separate 2 (Event Type) x 2 (Reaction Type) x 2 (Expertise) ANOVAs on judgments about PTSD behaviors, MDD behaviors, and distressed behaviors.³

**PTSD behaviors: Psychological abnormality and difficulty understanding judgments.** For psychological abnormality judgments of PTSD behaviors, there was a main effect of event type, such that PTSD behaviors were rated as less psychologically abnormal in the context of a traumatic event ($M = 5.10, SD = 2.12$) than in the context of an everyday event ($M = 6.60, SD = 3.08$).

³ DSM-5 experience was of less theoretical interest for these analyses than it was for PTSD diagnosis analyses. Prescribed judgments for psychological abnormality and difficulty understanding are not explicitly defined for disorders in the DSM-5 (APA, 2013), making any potential differences between DSM-5 experience for these dependent measures difficult to interpret. Therefore, we did not analyze the effect of DSM-5 experience on judgments of psychological abnormality and difficulty understanding.
1.60; F[1,128] = 22.88, p < .001, ηp² = .15). Please see Table B4. There was no main effect of reaction context (p = .788, ηp² = .01), no effect of expertise (p = .161, ηp² = .02), and no interactions (all p’s > .215, all ηp² < .02).

Also, for difficulty understanding judgments of PTSD behaviors, there was a main effect of event type, such that PTSD behaviors were rated as less difficult to understand in the context of a traumatic event (M = 2.50, SD = 1.18) than in the context of an everyday event (M = 4.99, SD = 1.98; F[1,128] = 78.12, p < .001, ηp² = .38). There was no main effect of reaction context (p = .886, ηp² < .01), or expertise (p = .704, ηp² < .01), and no interactions (all p’s > .130, all ηp² < .02).

To summarize, participants found PTSD behaviors to be less psychologically abnormal and less difficult to understand in traumatic event context compared to everyday event context, fully replicating past work demonstrating a proportionate-response effect (Kim et al., 2012a), wherein cases of life-events and behaviors that are proportionate in strength and valence (relative negativity; e.g., a traumatic event and PTSD behaviors) are rated as less psychologically abnormal and less difficult to understand compared to life-events and behaviors that are disproportionate in strength and valence (e.g., a everyday event and PTSD behaviors), which are rated as comparatively more psychologically abnormal and more difficult to understand.

**MDD behaviors: Psychological abnormality and difficulty understanding judgments.**

For psychological abnormality judgments of MDD behaviors, there was a main effect of event type, such that MDD behaviors, too, were rated as less psychologically abnormal in the context of a traumatic event (M = 4.76, SD = 2.10) than in the context of an everyday event (M = 6.75, SD = 1.45; F[1,52] = 40.30, p < .001, ηp² = .24). Please see Table B4. There was no main effect of reaction context (p = .486, ηp² < .01), or expertise (p = .410, ηp² < .01), and no interactions (all
Also, for difficulty understanding judgments of MDD behaviors, there was a main effect of event type, such that MDD behaviors were rated as less difficult to understand in the context of a traumatic event ($M = 2.53$, $SD = 1.26$) than in the context of an everyday event ($M = 5.13$, $SD = 2.14$; $F[1,128] = 70.97$, $p < .001$, $\eta_p^2 = .36$). There was no main effect of reaction context ($p = .627$, $\eta_p^2 < .01$), or expertise ($p = .595$, $\eta_p^2 < .01$), and no interactions (all $p$’s > .475, all $\eta_p^2 < .01$).

To summarize, participants found MDD behaviors to be less psychologically abnormal and less difficult to understand in traumatic event context compared to everyday event context, again cleanly replicating research demonstrating the proportionate-response effect (Kim et al., 2012a). Also of note is that participants were not influenced by event context when diagnosing depression for MDD, in line with its diagnostic criteria in the DSM-5 (APA, 2013), yet they were influenced by event context when making these other judgments about MDD behaviors. Taken together, these findings suggest that the influence of traumatic event context in considering MDD behaviors is dependent on the type of judgment being made.

**Distressed behaviors: Psychological abnormality and difficulty understanding judgments.** Finally, for psychological abnormality judgments of distressed behaviors, there was a main effect of event type, such that distressed behaviors were seen as more psychologically abnormal in the context of a traumatic event ($M = 2.71$, $SD = 1.39$) than in the context of an everyday event ($M = 2.20$, $SD = 1.17$; $F[1,128] = 4.21$, $p = .042$, $\eta_p^2 = .03$). Please see Table B4. There was no main effect of reaction context ($p = .711$, $\eta_p^2 < .01$) or expertise ($p = .544$, $\eta_p^2 < .01$), and no interactions (all $p$’s > .198, all $\eta_p^2 < .01$).

For difficulty understanding judgments of distressed behaviors, there was a main effect of event type, such that MDD behaviors were also rated as more difficult to understand in the
context of a traumatic event \((M = 3.03, SD = 1.82)\) than in the context of an everyday event \((M = 2.26, SD = 1.27; F[1,128] = 7.61, p = .007, \eta_p^2 = .06)\). There was no main effect of reaction context \((p = .455, \eta_p^2 < .01)\), or expertise \((p = .696, \eta_p^2 < .01)\), and no interactions (all \(p\)'s > .147, all \(\eta_p^2 < .01\)).

To summarize, participants found distressed behaviors to be more psychologically abnormal and more difficult to understand in traumatic event context compared to everyday event context, findings that are again consistent with the proportionate-response effect in that disproportionate context makes behaviors seem harder to understand and more abnormal (Kim et al., 2012a). Interestingly, we replicated this finding even though our vignettes differed from the everyday behaviors used in this past work, which found this pattern of judgments for a behavioral vignette that described someone exhibiting no change at all in behaviors after an event, and feeling unaffected by the event. In contrast, the current work employed a distressed behavior vignette describing someone exhibiting a non-disordered change in behaviors after the event, and feeling mildly affected by it.

**Experiment 1 implications and next steps.**

**Implications of the current findings for clinicians’ conceptualizations of PTSD as a DSM disorder.** In this study, clinicians’ diagnoses of PTSD aligned with the DSM-5 (APA, 2013) diagnostic criteria, and not with the DSM-IV-TR (APA, 2000) diagnostic criteria or research demonstrating a strong link between intense reactions and PTSD behavioral symptoms (e.g., Boals & Schuettler, 2009). Interestingly, clinicians and clinical trainees diagnosed PTSD more often in the same type of context that made PTSD behaviors seem less difficult to understand (i.e., traumatic event context), compared to the other types of context manipulated in this study (e.g., everyday event context). This strongly suggests that clinicians do not base PTSD
diagnoses upon intuitions about psychological abnormality. We explore this possibility in much more depth in a study reported in Appendix E and in the General Discussion section.

**Implications of the current findings for clinical training.** Clinicians and clinical trainees were differentially influenced by context when diagnosing PTSD for PTSD behaviors, MDD for MDD behaviors, and when noting that merely distressed behaviors were not disordered. Specifically, they were influenced by the contexts required for these diagnoses in the *DSM-5* (APA, 2013), in the different disorder-specific directions endorsed by that manual. Even clinical trainees, who had less experience in clinical practice than the clinicians in this study, made diagnostic judgments in accord with the most up-to-date diagnostic criteria for PTSD and MDD when diagnosing the disorders. Yet relative *DSM-5* experience did not impact diagnoses. About half of clinicians (58.93%) and trainees (45.0%) reported no or minimal *DSM-5* experience; furthermore, none of our analyses showed that self-reported *DSM-5* experience affected judgments. That is, at least half of participants likely could not have based their diagnoses on the *DSM-5* (APA, 2013), yet their diagnoses aligned with it nonetheless. We return to this finding in the General Discussion.

**Understanding why reaction information did not seem to influence judgments.** We uncovered no evidence that intense reaction context influenced the judgments of interest of either disordered behaviors (PTSD, MDD) or distressed behaviors. We also did not find that traumatic event and intense reaction contexts in conjunction influenced the judgments of interest. This finding was consistent across clinicians and clinical trainees (as well as lay people; see the first manuscript of this dissertation). It is noteworthy that people’s judgments aligned with the *DSM-5*’s (APA, 2013) criterion that PTSD may only be diagnosed in the context of a traumatic event, but did not align with research demonstrating that PTSD symptoms are highly likely to manifest
after a person experiences an intense reaction to any event (e.g. Boals & Schuettler, 2009).

However, we cannot conclude solely from null effects that people’s perception of the reaction information definitely does not influence people’s perceptions of PTSD. We believe it to be unlikely that people seemed uninfluenced by the reaction information simply because it was not as noticeable as event information in the text, as we were careful to ensure that event and reaction information were described equally saliently, equating them for length and interspersing both throughout the text of each vignette. Still, there are a number of plausible possible reasons why we found null effects for reaction context. One possibility is that people might have accurately encoded, stored and retrieved all parts of the vignette (i.e., event, reaction, and behaviors), but simply did not think that the reaction information was at all relevant to any of the dependent variables we assessed. For example, clinicians and clinical trainees might have inferred that reactions are not important because they are not included in the DSM-5 criteria for PTSD (APA, 2013). However, this possibility could not have applied to clinicians and trainees who had no experience with the DSM-5 (APA, 2013).

A second possibility is that people placed the greatest weight on the initially occurring factor (i.e., the event), as per prior research demonstrating a causal status effect (e.g., Ahn, 1998; Yopchick & Kim, 2009), and thus were not influenced by the secondary factor (i.e., the reaction to the event). However, some intermediate-strength influence of the secondary factor would still have been expected if this were the case; yet we found no evidence that supported this possibility. A third possibility is that people might not tend to remember reaction information accurately, and instead retrieve a distorted memory of the reaction (Brainerd & Reyna, 2002) that is more proportionate to the event or behaviors than the actual description of the reaction. In Experiment 2, we examined this final possibility: whether clinicians are able to recognize event,
reaction, and behavior information (from the vignettes we used in Experiment 1) with equal accuracy. We were particularly interested in whether there are any differences in clinicians’ memories for reactions, as compared to events and behaviors. (The first two possibilities were outside the scope of the current investigation, but please see the General Discussion for further consideration of these issues.)

**Experiment 2: Clinicians’ and Clinical Trainees’ Memories of Events, Reactions, and Behaviors**

In Experiment 2, we investigated whether clinicians and clinical trainees are particularly susceptible to creating false memories of reaction context, compared to event context and behaviors. We used three vignettes from Experiment 1, including the same event and reaction contexts. After clinicians and clinical trainees were asked to carefully read and understand the vignettes and complete a short filler task, we presented them with a recognition task containing original (correct) elements of the events, reactions, and behaviors, false elements that were matched to the original ones in terms of thematic content but not strength (e.g., mild versus severe), and false elements that were matched to the original ones in strength but not thematic content. Our analyses were aimed at understanding whether clinicians and clinical trainees falsely recognized reaction information more often than event and behavior information. We were also especially interested in whether participants’ memories of reaction context were systematically distorted in strength to match the events and/or behaviors.

Experiment 2 investigated Questions 6 and 7. Question 6 asked whether clinicians formed distorted memories for reactions more often than for events or behaviors. Question 7 asked whether participants were particularly susceptible to distorting their memory of the
reaction information such that it matched the strength of negative valence of the event or behavior information. Please see the Introduction (p. 77) for hypotheses.

Questions 1, 2, 3, and 5: Replications

Finally, we aimed to examine whether our findings for questions 1, 2, 3, and 5 would be replicated in a second, separate sample of clinicians and clinical trainees in Experiment 2. Past work in the cognitive literature had not yet explored whether open-ended diagnostic judgments are influenced by context. Thus, it was important to investigate whether our Experiment 1 findings for open-ended diagnoses could be replicated in Experiment 2. We did not re-examine Question 4; because the memory task lengthened the procedure considerably, and because we already fully replicated the proportionate-response effect in abnormality and difficulty understanding judgments in Experiment 1 from past work, these measures were not included in Experiment 2.

Methods

Participants.

Clinicians. Thirty-four licensed practicing clinicians (20 female) participated in response to a mailed postcard invitation to an online study. Clinicians who completed the study were sent a $25 gift certificate to an online retailer via email. To recruit these clinicians, a random sample of 200 clinicians from 47 states in the U.S. was taken in the same way as in Experiment 1. Of the 200 clinicians who were mailed postcards, four clinicians’ cards were returned because the clinician had moved. This yielded a 17.0% response rate (comparable to Experiment 1 and to past work; e.g., Rottman, Ahn, Sanislow, & Kim, 2009).

Clinicians reported a mean age of 47 years (range: 24-78), and most self-identified as White (94%); 3% self-identified as Black, and the remaining participants declined to report race.
In addition, 3% self-identified as Hispanic. The majority of clinicians held a private practice (74%), but some also practiced in other settings (6% hospital, 6% community clinic, 3% nursing home, 3% family medicine, 6% psychology department clinic, 2% declined to report). Clinicians had a median of 16.5 years of clinical experience (interquartile range: 6 years, 29.5 years) and held a range of primary clinical orientations (38% Cognitive / Cognitive-Behavioral, 18% Psychoanalytic, 6% Humanistic, 9% Eclectic, 9% Family Systems, 3% Behavioral, 14% other, and 3% declined to report).

**Clinical trainees.** Thirty-seven graduate students (29 female) were also recruited in the same way as Experiment 1. Those who completed the study were entered into one of several raffles for a $25 gift card to an online retailer. Graduate students reported a mean age of 27 years (range: 22 - 45), and self-identified as White (84%), Asian (8%), and Black (5%; the remaining 3% declined to report). In addition, 5% self-identified as Hispanic. They had a median of 2 years of clinical experience (interquartile range: .5 years, 2 years) and provided clinical services in a variety of settings (59% psychology department clinics, 11% hospitals, 5% community clinics, 14% other, 11% had not yet provided clinical services). Graduate students’ primary clinical orientation was most frequently Cognitive / Cognitive-Behavioral (49%), followed by Behavioral (14%), Eclectic (14%), and others.

**Materials.**

**Recognition items.** The recognition items consisted of correct items describing true information from the original vignettes and incorrect items describing false information from the vignettes (detailed information on norming criteria is provided in the next section). Correct and filler items were paraphrased from the original vignette, rather than repeated verbatim, to eliminate the possibility that clinicians and clinical trainees might merely recognize specific
combinations of words from the original vignettes in the main study (this could produce ceiling effects). For example, if the original vignette read, “A tractor-trailer came out of nowhere and slammed into his car,” a correct item might be, “Martin was hit by a tractor-trailer.” This is intended to be true of the statement, though the wording has been paraphrased (see norming ratings in the next section).

All incorrect items were designed to describe reasonably plausible occurrences given the original vignette, as highly implausible material would be too easy to reject, and might produce floor effects in the main study. There were two types of incorrect items. Incorrect content items described information not in the original vignette, but that was just as severely or mildly negative as the information provided in the original vignette. An incorrect content item might state, “Fire trucks arrived at the scene.” This item would be false, as it did not occur in the original vignette, although it is plausible that something like this could happen under circumstances such as those described in the vignette, and is written to be equivalently severe as the original statement (again, see norming data below). Incorrect strength items took real information from the original vignette, but described it as either occurring much more severely or much more mildly negatively than in the original vignette. An example of an incorrect strength item might be, “The tractor-trailer merely tapped Martin’s car.” This item would also be false, but this time because it is much weaker than the original statement. Manipulating the incorrect strength items as such allowed us to test our question about whether clinicians’ and clinical trainees’ memories of reaction information is distorted in strength, in particular, to align with the strength of the event or behavior information.

For Experiment 2, we created recognition items for 3 hypothetical people. For each hypothetical person, we created recognition items for seven vignette elements: Two event
elements (traumatic and everyday), two reaction elements (intense and mild), and three behavior elements (PTSD, MDD, and distressed). Given that there were three hypothetical people and seven vignette elements per hypothetical person, there were 21 vignette elements total. For each vignette element, we created eight recognition items: four true (correct) items, two false content items, and two false strength items. We needed to employ twice as many “true” items (4) as false content items (2) or false strength items (2), so that there were an equivalent number of “true” items (4) and “false” items (4) overall, which allowed us to perform d’ analyses. Given that there were 8 recognition items per vignette element, and 21 vignette elements total, we created 168 recognition items total. Please see Tables B5-B7 for all recognition items. The recognition items were normed in the following three norming experiments.

**Recognition items: True / false norming.** To be used as stimuli in the main study, the correct items had to be rated by people as true of the vignette and the incorrect items had to be rated by people as false (not true of the vignette). Thus, the first step of this norming procedure consisted of eliciting “true” or “false” responses from people who were asked to rate whether each of the 168 items accurately described the original vignette. The vignette was always displayed on the same page as the recognition item so that participants could directly compare whether the two were consistent if they wished, and participants always rated one item per page by clicking a button on the screen indicating either “true” or “false.”

In all three true/false norming experiments, participants rated all eight recognition items for a single vignette (e.g., Martin’s traumatic event) before answering the next eight recognition items for the next vignette, until all recognition items for all vignettes had been rated. This was done to cut down on overall experiment time because participants could keep one vignette in mind, rather than being asked to repeatedly switch between vignettes. Vignettes were also
always presented in randomized order across the norming experiments. None of the participants who participated in the norming studies took part in the main study).

In the first true/false norming experiment, seven undergraduate students participated, unanimously agreeing on 150 of the 168 items, for which they rated all original correct items as “true” and all incorrect false items as “false.” Next, the 18 items that were not rated unanimously as either true or false, depending on the item, were edited by the experimenters. Then, we again elicited “true” or “false” responses from the same set of seven students, who unanimously rated all remaining true items as “true” and all remaining false items as “false” in the second round.

In the second true/false norming experiment, we elicited true / false ratings from 13 lay participants recruited from Amazon Mechanical Turk. They saw all 168 items (as edited in the first norming experiment) and unanimously rated 138 of the 168, indicating that the original items were “true” and the false items were “false.” The remaining 30 items went through a round of edits by the experimenters. Then, a separate set of 12 participants recruited from Amazon Mechanical Turk rated the 30 updated items, unanimously rating all but five. These five items were edited and included in the updated set of 136, which we next tested on participants with extensive clinical experience. It was necessary to elicit judgments from people with a background in clinical psychology because this population constituted our sample in Experiment 2.

Finally, in the third true/false norming experiment, we elicited true / false ratings from 15 participants with extensive training in clinical psychology, recruited by email. All indicated that they were either a psychology intern, a postdoctoral fellow in clinical psychology, or in their fifth or sixth year of training. These participants unanimously rated 161 out of 168 recognition items as “true” for true and “false” for false. We edited the remaining seven items and recruited
another set of five participants from the same recruitment pool to judge the seven updated items. They agreed on all but one item, which was updated and tested on a final set of three clinical psychology participants, who rated it unanimously.

**Recognition items: Strength norming.** We also needed to ensure that the false strength items truly did mismatch the original items in strength of negative valence. Specifically, we wanted to check that both the original correct items and the false content items differed in severity from the false strength items, and that the original correct items and the false content items were equivalently severe. Nine participants from Amazon Mechanical Turk rated each recognition item on the same severity scale used for norming the stimuli in Experiment 1. That is, they rated each recognition item on a scale of 1-9, where 1 = not at all severe, and 9 = extremely severe. In this norming study, the vignette was not presented on the same page as the items, as in the other norming studies, because it was irrelevant to the task. Because we were interested in general severity, we included the severe items together (i.e., traumatic events, severe reactions, PTSD behaviors, MDD behaviors) and the mild items together (i.e., everyday events, mild reactions, distressed behaviors) for analyses.

To examine whether false strength items differed in severity from correct items and false content items, we ran two separate repeated-measures ANOVAs with recognition item type (correct original, false content, false strength) as the factor, one for severe items and one for mild items. For severe items, there was a main effect of recognition item type ($F[2,16] = 144.42, p < .001$). The main effect also appeared for mild items ($F[2,16] = 91.75, p < .001$). Bonferroni-corrected paired samples t-tests revealed that the severe false strength items (which were designed to be mild; $M = 3.28, SD = 0.87$) were rated as less severe than the severe original correct items ($M = 6.66, SD = 1.11; t[8] = 14.62, p < .001$) and also as less severe than the severe
false content items ($M = 6.75, SD = 1.07; t[8] = 11.57, p < .001$). Also, the mild false strength items (which were designed to be severe; $M = 7.30, SD = 1.09$) were rated as more severe than the mild original correct items ($M = 3.85, SD = .90; t[8] = -9.75, p < .001$) and as more severe than the mild false content items ($M = 3.75, SD = .78; t[8] = -11.35, p < .001$). Please see Figure B4. Thus, the false strength items did differ in severity from the vignette, fulfilling the desired severity criteria. Also of interest, both original correct items were rated as equivalently severe as both false content items (severe items: $t[8] = -.65, p = .533$; mild items: $t[8] = .46, p = .661$).

**Recognition items: Content norming.** We needed to ensure that the false content items truly did mismatch the vignette in content. We also wanted to know that the false strength items matched the vignette in content to the same extent that the original correct items did so. To test this, we recruited nine participants from Amazon Mechanical Turk, who rated each item (referred to as a “statement” in this norming study) on a scale of 1-9, where 1 = this statement describes the vignette very poorly, and 9 = this statement describes the vignette very well. Each recognition item was presented on a separate page, with the vignette always printed above it for reference.

Most importantly, the false strength items ($M = 3.03, SD = 1.87$) were rated as equivalently good descriptions of the vignette as the original correct items ($t[8] = 2.08, p = .071$), as they were designed to be. In addition, the false strength items were not judged to be better descriptions of the vignette than the false content items ($t[8] = -2.41, p = .042$), which was also expected, as the false strength items were designed to be imperfect descriptions of the original vignette, describing some content from them, but more weakly or strongly. To examine whether false content items were judged to differ in content from the vignettes, we asked whether they differed more so than the original correct items and more so than the false strength items. We
conducted a repeated-measures ANOVAs with recognition item type (correct original, false content, false strength) as the factor, which yielded a significant main effect \( F[2,16] = 6.76, p = .007 \). Bonferroni-corrected paired samples t-tests were run to examine this effect, and revealed that the false content items \( M = 2.61, SD = 1.98 \) were rated as worse descriptions of the vignette than the original correct items \( M = 4.01, SD = 1.59; t[8] = 3.07, p = .015 \), thus fulfilling this desired criterion.

Overall, the norming studies confirmed that the recognition items adhered to the given criteria. In pretesting, research assistants, lay participants, and people with extensive background in clinical psychology rated the correct items as “true” and the false items as “false.” The original correct items and the false content items differed in severity from the false strength items, and the original correct items and the false content items were equivalently severe. Finally, the false strength items and the original correct items were perceived as equivalently good descriptions of the vignette in terms of content, and the false content items were confirmed to mismatch the correct items in content.

**Procedure.** In the main study, clinicians and clinical trainees read three vignettes from Experiment 1, each describing a different person with PTSD behaviors, MDD behaviors, and distressed behaviors, respectively. Each vignette described one of the three hypothetical people (i.e., “Martin,” “Cathy,” or “Justin”). The behaviors were counterbalanced across the three hypothetical people, such that one third of participants read about Martin with PTSD behaviors, Cathy with MDD behaviors, and Justin with distressed behaviors. Thus, another third of participants read about Cathy with PTSD behaviors, Justin with MDD behaviors, and Martin with distressed behaviors. The final third of participants read about Justin with PTSD behaviors, Martin with MDD behaviors, and Cathy with distressed behaviors. Each participant saw all three
behavioral vignettes presented in one of four contexts as determined by the 2 (Event Type: Traumatic, Everyday) x 2 (Reaction Type: Severe, Mild) between-subjects design from Experiment 1. Participants were also sorted randomly into one of three combinations of behaviors counterbalanced across hypothetical people.

**Diagnosis task.** For each vignette, we asked clinicians and clinical trainees to complete the exact same diagnosis task as in Experiment 1. The diagnosis task also acted as an encoding task in which participants had to read and understand each vignette.

**Filler task.** The diagnosis task was followed by a five-minute filler task that was unrelated to the content of the vignettes. As is standard in memory research, the purpose of the task was to reduce the likelihood that clinicians and clinical trainees are able to maintain the vignette information in working memory, thereby reducing the likelihood of ceiling effects in their recognition task performance. Specifically, we showed them Rorschach ink blots and other ambiguous figures in the filler task, and asked clinicians and clinical trainees to report what they saw, in addition to requesting answers to GRE-level math problems. There were 60 filler questions in total, presented in randomized order. The filler task terminated after five minutes, regardless of how many questions had been answered. No one completed the entire available filler task, and no participants chose to exit the study during the filler task.

**Recognition task.** Finally, participants were presented with a recognition task in which they were asked to indicate whether they recognized information from the original vignettes. The dependent variable was a response of “true” or “false.” Participants saw each true / false question on the same page as one recognition question.

For example, for the vignette about “Martin,” they were asked eight questions about the event. Specifically, they were asked to indicate whether they recognized (1) two pieces of correct
information from the original vignette, (2) two pieces of false information that were of the same strength as in the original vignette but differed in content (*false content* information), (3) two pieces of false information that had a different strength from the original vignette but described content from the vignette equivalently well (*false strength* information), and (4) two other pieces of correct control information from the original vignette. Again, please see Tables B5-B7 for all recognition items.

In total, participants made 72 recognition judgments. Specifically, they judged eight items for each of the three events that described different hypothetical people (24 event items), eight items for each of the three reactions describing those hypothetical people (24 reaction items), and eight items for each of the three sets of behaviors, one for each hypothetical person (24 behaviors items).

We randomized the order of recognition questions such that participants answered all 24 recognition questions about a single vignette (e.g., Martin), in randomized order, before answering all recognition questions about the next vignette, followed by the third vignette.

Participants only judged a subset of the 168 items normed. This was because they learned about only one of the two types of events (traumatic events or everyday events), only one of the two types of reactions (intense reactions or mild reactions), and only one of the three combinations of PTSD, MDD, and distressed behaviors counterbalanced across the three hypothetical people (please see procedure). Overall, each of the two event elements were judged by 35 or 36 people, each of the two reaction elements were judged by 35 or 36 people, and each of the three behaviors elements were judged by 23 or 24 people.
We recorded both people’s responses to the recognition items and how long they took to answer each question. Demographic questions and a debriefing statement concluded the study. The entire main experiment took approximately 40-60 minutes to complete.

**Results and Discussion**

**Open-ended diagnoses analyses.** Because participants made diagnosis judgments in the first task, we were also able to examine whether our results from Experiment 1 on this task were replicated. For the encoding task, where people made open-ended diagnoses, we again looked at PTSD diagnoses for PTSD behaviors, depression diagnoses for MDD behaviors, and “none” diagnoses for distressed behaviors using chi-square tests of event context, reaction context, and expertise. Again, PTSD behaviors were diagnosed more often as PTSD in traumatic event context (94%) than in everyday event context (66%; $X^2[1, N = 71] = 9.25, p = .002$; see Figure B3). For PTSD behaviors, the tests of reaction context and expertise were non-significant (all $p$s $>.170$). This replicates the Experiment 1 finding (p. 89) that traumatic event context in particular led participants to diagnose PTSD for PTSD behaviors.

The chi-square tests examining the influence of event and reaction context on depression diagnoses for MDD behaviors were non-significant (all $p$s $>.192$; see Figure B3 for depression diagnoses for MDD behaviors by event context). This replicates the Experiment 1 finding (p. 92), that participants did not rely on context to diagnose depression for MDD behaviors. However, unlike in Experiment 1, the test of expertise was significant, such that clinical trainees made more depression diagnoses for MDD behaviors (73%) than clinicians did (44%; $X^2[1, N = 71] = 6.11, p = .013$). In addition to depression, our practicing clinician participants diagnosed adjustment disorder with depressive features (21%), PTSD (15%), sub-clinical depression (6%), and others (14%), and rated those diagnoses as their most likely diagnosis for the case of MDD.
behaviors. Although we took care to ensure that the MDD vignettes did not qualify for a *DSM* diagnosis of adjustment disorder (by describing symptoms as beginning or continuing more than three months past an experienced stressor, the cut-off point for adjustment disorder), clinicians may have been particularly swayed to diagnose adjustment disorder and PTSD to acknowledge the event that the person was responding to, even if it was merely an everyday event.

For distressed behaviors, we found that indications of “none” were uninfluenced by event context, reaction context, or expertise (all *p* > .351; see Figure B3 for “none” diagnoses for distressed behaviors by event context). This diverges from the Experiment 1 finding (p. 92), wherein participants made fewer “none” diagnoses in traumatic event context than in everyday event context, and instead diagnosed a variety of disorders in traumatic event context. This may be because in Experiment 2 we only asked participants to diagnose three vignettes, compared to six in Experiment 1. Perhaps with only three vignettes total, compared to six in the previous study, the distressed behaviors contrasted sufficiently with the two disordered behaviors such that participants viewed the distressed behaviors as clearly non-disordered.

**Recognition task analyses.** In these analyses, the primary goal was to examine whether clinicians and clinical trainees are more likely to make more false alarms for reactions than events or behaviors (Question 6), and also whether they are more likely to have formed systematically distorted memories for reactions whose strength is disproportionate to the events or behaviors (Question 7). We also addressed the question of whether people’s performance was affected by expertise, and whether we replicated our results from Experiment 1 on the open-ended diagnosis task (Questions 1, 2, 3, and 5).

**False alarms for reactions.** To address Question 6, whether clinicians and clinical trainees made more false alarms for reactions than for events or behaviors, we performed a
repeated measures ANOVA on the proportion of false alarms, with item type (events, reactions, and behaviors) entered as the factor. The analysis showed that the mean proportion of false alarms differed significantly between item types ($F[2, 140] = 32.25, p < .001$). Bonferroni-corrected paired samples t-tests revealed that clinicians and clinical trainees made a greater proportion of false alarms for reactions ($M = .28, SD = .16$) than for events ($M = .18, SD = .15$; $t[70] = 5.22, p < .001$) or for behaviors ($M = .13, SD = .12$; $t[70] = 7.70, p < .001$). Please see Figure B5. This finding lends some support to our conjecture that clinicians and clinical trainees were less influenced by reaction information in Experiment 1 because they did not remember the reaction information as accurately as they did the event or behavior information. Interestingly, participants also made fewer false alarms for behaviors than events ($t[70] = 2.88, p = .005$). This may be in part because the encoding task asked clinicians and clinical trainees to make diagnoses; most DSM diagnoses require an assessment of behavioral symptoms only, not events or reactions, and clinicians and clinical trainees might be generally more attentive to them.

**Hits for reactions.** To distinguish whether clinicians and clinical trainees indeed had a tendency to form more distorted memories for reactions (indicated by more false alarms), or whether they simply used a very low response criterion for reactions, responding “yes” to almost everything (which would be indicated by both more false alarms and more hits), we ran a repeated measures ANOVA on the proportion of hits, with event, reaction, and behaviors as factors. The ANOVA showed that the mean proportion of hits differed marginally between item types ($F[2, 140] = 2.79, p = .065$); however, Bonferroni-corrected paired samples t-tests showed that clinicians and clinical trainees had an equivalent proportion of hits for events ($M = .87, SD = .12$), reactions ($M = .84, SD = .12$), and behaviors ($M = .82, SD = .15$; all $p$s $\geq .039$). Please see Figure B6. Also, a one-sample t-test indicated that the proportion of hits for events ($t[70] = -$
8.98, \( p < .001 \), reactions (\( t[70] = -10.91, p < .001 \)), and behaviors (\( t[70] = -10.16, p < .001 \)) differed from ceiling (i.e., a proportion of 1.0, indicating 100% hits). These results show that clinicians and clinical trainees did not have a higher proportion of hits for reactions than for events or behaviors. Considering the above analyses for false alarms and hits in conjunction, therefore, we suggest that clinicians and clinical trainees were indeed particularly susceptible to forming distorted memories for reactions, compared to events and behaviors.

*False alarms and hits for reactions: Effects of event and reaction disproportionality.*

We next sought to determine whether clinicians’ and clinical trainees’ memories for reactions were distorted in strength of negative valence to become more proportionate to the events or behaviors (Question 7). To do so, we analyzed the proportion of false alarms for false strength reaction questions in conditions where the event and/or behaviors were proportionate to the reaction, and in conditions where the event and/or behaviors were disproportionate to the reaction. First, we compared the conditions where the events were proportionate to the reactions against the conditions where the events were disproportionate to the reactions. Next, we compared the conditions where the reactions were proportionate to the behaviors against the conditions where the reactions were disproportionate to the behaviors.

We coded the following event and reaction conditions as proportionate, collapsing across behavior types: The traumatic event / intense reaction condition (both severely negative) and the everyday event / mild reaction condition (both mildly negative). We also coded the following as disproportionate conditions: The traumatic event / mild reaction condition (severely negative / mildly negative) and the everyday event / severe reaction condition (mildly negative / severely negative). Because these were between-subjects factors, we were able to run a single independent samples t-test, to compare the proportion of false strength false alarms for reactions
in disproportionate event context to that in disproportionate event context. The paired samples t-test revealed that people made a higher proportion of false strength false alarms for reactions in disproportionate conditions ($M = .33$, $SD = .17$) than in proportionate conditions ($M = .24$, $SD = .19$; $t[69] = -2.06$, $p = .043$).

The above false strength questions consisted of information from the original vignette that was distorted to be weaker or stronger in negative valence. This means that people incorrectly recognized reaction information that was distorted in strength to match the event, more so than they incorrectly recognized reaction information that was distorted in strength to mismatch the event. In other words, clinicians and clinical trainees were indeed more susceptible to forming distorted memories of reactions that better matched the strength of the event (Question 7).

*False content items.* To determine whether the above effects held true only for false strength items, we ran the above independent samples t-test for false alarms for false content items. A t-test found no significant difference between the two ($t[69] = -.60$, $p = .548$), indicating that false content false alarms did not occur more often in disproportionate vs. proportionate event and reaction conditions. Thus, the effect of the event and reaction proportionality manipulation on false alarms held true only for false strength items, and did not affect people’s propensity to falsely remember incorrect content. Note, however, that people did have false alarms for incorrect content overall, in keeping with past classic work (Deese, 1959; Roediger & McDermott, 1995): The mean proportion of false content false alarms was $.28$ ($SD = .22$), suggesting that people falsely recognized the false content about one quarter of the time.

Finally, we ran the same independent samples t-test for hits for reactions, to confirm whether hits were similarly unaffected by the proportionality between events and reactions. This
test was also non-significant ($t[69] = -0.81, p = 0.420$), indicating that the event and reaction manipulation did not affect participants’ ability to correctly recognize reactions from the original vignette.

**False alarms and hits for reactions: Effects of reaction and behavior disproportionality.** To continue addressing Question 7, we wanted to see whether people were also susceptible to forming distorted memories of reactions so that they matched the behaviors in strength. To do so, we coded the PTSD and MDD behavioral vignettes as severe behaviors and coded the distressed behavioral vignette as mild behaviors. Thus, the proportionate conditions were severe reaction / severe behaviors and mild reaction / mild behaviors and the disproportionate conditions were severe reaction / mild behaviors and mild reaction / severe behaviors. Because the behavior type conditions were within-subjects, we ran two separate paired samples t-tests, each comparing severe behaviors to mild behaviors, one in severe reaction context and one in mild reaction context. In each test, the dependent variable was the proportion of false alarms for false strength reaction items. This time we were interested in the match or mismatch between reactions and behaviors, so the data were collapsed across event type. For severe reactions, a paired samples t-test showed that people made more false alarms in the disproportionate severe reaction / mild behaviors condition ($M = 0.46, SD = 0.31$) than in the proportionate severe reaction / severe behaviors condition ($M = 0.20, SD = 0.19$; $t[34] = -4.54, p < 0.001$). For mild reactions, the paired samples t-test was also significant ($t[35] = 2.51, p = 0.017$), such that people made more false alarms in the disproportionate mild reaction / severe behaviors condition ($M = 0.33, SD = 0.24$) than in the proportionate mild reaction / mild behaviors condition ($M = 0.18, SD = 0.30$). See Figure B7.
Thus, clinicians and clinical trainees made more false strength false alarms when the reaction and behaviors described in the vignette were disproportionate in strength than when the event and behaviors matched in strength. The false strength items in the recognition task presented reaction information from the original vignette more mildly or strongly than in the original vignette, meaning that clinicians and clinical trainees who endorsed false strength items in disproportionate behavior conditions falsely remembered the reactions as aligning in strength with the behaviors. Participants may have seen the reaction as not merely a precursor to the behaviors but as a short, early sign of them, leading them to group the two together into one memory where they remembered all information as consistent. We speculate that clinicians and clinical trainees may be for the most part encoding and directly retrieving information about behaviors (and events); our first set of analyses shows that they are less likely to encode and retrieve reaction information than the other two types, and the proportionality analyses just reviewed suggest that clinicians and clinical trainees may instead infer the reaction information from the behaviors and/or events.

To check whether the influence of reaction-behavior proportionality on false alarms was specific to false strength items, we also ran the above analyses on false alarms for false content reaction items. For severe reactions, the paired samples t-test was not significant ($t[34] = 1.38, p = .176$), nor was it significant for mild reactions ($t[35] = -.16, p = .872$). That is, people made false alarms for false content items without being influenced by the proportionality between reactions and behaviors. Finally, we ran these analyses again for hits for reaction questions, to check that clinicians’ and clinical trainees’ ability to correctly recognize reactions from the original vignette was not impeded or facilitated by a match or mismatch between reactions and
behaviors. For hits, again, we found that neither the paired samples t-test for severe reactions ($t[34] = 1.35, p = .186$) nor the t-test for mild reactions was significant ($t[35] = 1.39, p = .172$).

Overall, these findings strongly suggest that clinicians’ and clinical trainees’ memories of the reaction were brought into alignment with the event and behavior information, such that the reaction strength retrieved was altered (Question 7). We speculate this may have occurred because clinicians and clinical trainees’ only remembered the most basic information about the reactions, that the reactions occurred at all, and then inferred the strength of the reaction information as proportionate to the event and behaviors. That said, the current research cannot distinguish whether this systematic distortion occurred during encoding, storage, or retrieval, and this remains an open question for further investigation (please see also General Discussion).

**D prime.** To further address Question 6, we determined whether participants were particularly insensitive to the difference between true and false reactions. As a measure of sensitivity, $d'$ was calculated for events, reactions, and behaviors. Forty-one percent of participants’ responses for events, 16% for reactions, and 38% for behaviors were perfect, with participants either making all possible hits or not committing any false alarms. These responses were converted using a standard correction method (Macmillan & Kaplan, 1985). A repeated-measures ANOVA with vignette element (event, reaction, behaviors) as the factor showed that there were differences between the item means ($F[2,140] = 16.77, p < .001$). Bonferroni-corrected paired samples t-tests were run to further investigate the effect. Participants showed a greater sensitivity to events ($d' = 2.18$) than to reactions ($d' = 1.72; t[70] = 5.21, p < .001$) and also a greater sensitivity to behaviors ($d' = 2.23$) than to reactions ($t[70] = -5.32, p < .001$). They showed an equivalent sensitivity to events and behaviors ($t[70] = -.49, p = .628$). Thus, clinicians
and clinical trainees were better able to distinguish between true and false events and between true and false behaviors than they were able to distinguish between true and false reactions.

Next, we were interested in whether participants demonstrated particular insensitivity to reactions that were disproportionate to events. We again ran a repeated-measures ANOVA with vignette element (event, reaction, behaviors) as the within-subjects factor, adding event / reaction proportionality as the between-subjects factor (please see bottom of p. 55 for details on how this was coded). The test demonstrated neither a main effect of event / reaction proportionality nor an interaction of event / reaction proportionality and vignette element (all \( p \geq .339 \)). Thus, the proportionality between events and reactions did not influence participants’ sensitivity.

Finally, we investigated whether participants’ insensitivity was influenced by the proportionality between reactions and behaviors. To do so, we calculated \( d' \) for events, reactions, and behaviors that were presented in vignettes with severe behaviors (PTSD and MDD behaviors) and mild behaviors (distressed behaviors). The following percentages of responses were perfect, therefore converted with the standard correction method (Macmillan & Kaplan, 1985): 60% of events, 35% of reactions, and 58% of behaviors. We performed two repeated-measures ANOVAs, one for vignette items presented in vignettes with severe behaviors, one for mild behaviors vignette items, with vignette element (event, reaction, behaviors) as the within subjects factor and reaction condition (intense or mild) as the between subjects factor. For the items presented in vignettes with severe behaviors, the ANOVA showed neither a main effect of reaction context nor an interaction of vignette item and reaction context (all \( p \geq .417 \)). However, for the items in mild behavior vignettes, the ANOVA revealed a main effect of reaction context (\( F[1,69] = 5.60, p = .021 \)), such that people had better sensitivity to recognition items presented in vignettes when the reaction and behaviors were proportionate (mild reaction and mild
behaviors; \(d' = 1.63\) than when they were disproportionate (intense reaction and mild behaviors; \(d' = 1.40\)). However, there was not an interaction \(F[2,138] = .25, p < .780\). This second test shows that when an intense reaction leads to disproportionately mild behaviors, people are prone to demonstrating a worse memory for information from such a vignette, overall.

We also compared the sensitivity of clinicians and clinical trainees on the recognition task overall (Question 5). To do so, we calculated \(d'\) for each group across all recognition judgments. Four percent of participants’ responses were perfect, so we converted those responses using the same standard correction method (Macmillan & Kaplan, 1985). An independent-samples \(t\)-test was conducted to compare \(d'\) for clinicians and trainees, and showed that clinical trainees had higher sensitivity \((M = 2.24, SD = .58)\) than clinicians \((M = 1.79, SD = .49; t[69] = -3.52, p = .001)\). That is, clinical trainees were better able to distinguish accurate information presented in the original vignette from false information than clinicians. To further investigate this finding, correlational analyses were conducted to examine whether years of clinical experience predicted sensitivity, and we found this to be the case. Years of clinical experience and \(d'\) were significantly negatively correlated \((r = -.301, p = .011)\). That is, the more years of clinical experience a participant had, the more difficult it was for them to distinguish correct information from false lures. Because we found a similar effect with reaction times, please see the section below for further discussion of this finding.

**Reaction times.** In order to examine the effects of vignette element (event, reaction, behaviors) on reaction time, we conducted three repeated measures ANOVAs with vignette element as the factor (event, reaction, behaviors) and mean reaction time as the dependent variable. Separate ANOVAs were conducted for hits, false strength false alarms, and false content false alarms. All responses over 15000 ms were trimmed before analyses to avoid
including data most likely skewed by uncontrolled events (e.g., sneezing); no responses were under 300 ms (Reis & Judd, 2014)\(^4\). For hits, overall RT for events \((M = 2334 \text{ ms}, SD = 774)\), reactions \((M = 2250 \text{ ms}, SD = 762)\), and behaviors \((M = 2334 \text{ ms}, SD = 970)\) did not significantly differ \((F[2,140] = .65, p = .522)\). For false strength false alarms, overall RT for events \((M = 2668 \text{ ms}, SD = 941)\), reactions \((M = 2819 \text{ ms}, SD = 340)\), and behaviors \((M = 2558 \text{ ms}, SD = 1070)\) also did not differ \((F[2,140] = .31, p = .738)\). Additionally, for false content false alarms, overall RT for events \((M = 2543 \text{ ms}, SD = 922)\), reactions \((M = 2690 \text{ ms}, SD = 1233)\), and behaviors \((M = 2481 \text{ ms}, SD = 1066)\) did not differ \((F[2,140] = 1.26, p = .288)\). Thus, there was no indication that participants took more time processing the recognition questions for one type of vignette element. Also, it suggests that participants when fared worst on the reaction recognition items in terms of accuracy, this was not because they spent less time processing them. Because participants did not take a different amount of time to answer reaction questions compared to event or behavior questions, we did not further examine whether the proportionality between reactions and events, or between reactions and behaviors, also affected reaction times.

To compare the reaction times of clinicians to those of clinical trainees, we ran a univariate ANOVA on each participant’s average reaction time, with expertise as the factor. The analysis revealed that clinicians took longer to respond to recognition questions \((M = 2782 \text{ ms}, SD = 760)\) than clinical trainees \((M = 2145 \text{ ms}, SD = 593; F[1,69] = 15.61, p < .001)\). To further investigate, correlational analyses were conducted to examine whether years of clinical experience predicted reaction times. Indeed, years of clinical experience and reaction time were positively significantly correlated \((r = .532, p < .001)\). That is, the longer participants had been practicing, the longer they took to make decisions. In addition, as described in the preceding

\(^4\) Reaction time data should be interpreted with caution as these data were collected online.
section, we found that clinicians had worse sensitivity than clinical trainees, which was correlated with years of clinical experience. Although these findings might seem surprising at first glance, it stands to reason that more experienced clinicians have learned about a great many clients’ behavioral symptoms over time, and therefore might find it more difficult than relative novices (who have much less information about clients stored in memory) to precisely map any given set of features to a particular case.

**General Discussion**

In this work, we addressed seven research questions; the first two were central and the last five followed from these. The first question we asked was whether clinicians and clinical trainees’ diagnoses of PTSD are influenced by the specific type of context required for a PTSD diagnosis in the *DSM-5* (i.e., traumatic event context only; APA, 2013), the *DSM-IV-TR* (i.e., both traumatic event and intense reaction contexts; APA, 2000), or neither. In Experiment 1, participants diagnosed PTSD more consistently in the context of a traumatic event than in the context of an everyday event, in line with the *DSM-5* (APA, 2013) criteria, rather than only in the context of both a traumatic event and an intense reaction. In Experiment 2, we replicated this finding.

Our second question was whether context increases or decreases PTSD diagnoses. In both Experiments 1 and 2, proportionately negative traumatic event context increased PTSD diagnoses relative to disproportionately mildly negative everyday event context.

Third, we asked whether the effects of context on PTSD diagnoses could be generalized to another disorder (i.e., MDD) and diagnostic judgments about everyday behaviors. We did not find this to be the case for MDD. MDD behaviors were diagnosed with depression regardless of event context, demonstrating that context’s influence on diagnoses may in fact be disorder-
specific. One likely possibility is that PTSD was diagnosed more often in traumatic event context because it has long been included in the very definition of PTSD (i.e., DSM diagnostic criteria). Indeed, the diagnostic criteria for MDD at present do not include any context, and we found that MDD was diagnosed in this way. Both disorder diagnoses adhered to the DSM-5 (APA, 2013) criteria in this study (even though, again, many had little to no familiarity with this version of the manual). More work is needed to examine whether diagnoses for other disordered behavioral symptoms are increased in traumatic event context and whether they also adhere to the DSM-5 (APA, 2013).

To examine whether traumatic event context increased diagnostic judgments about behaviors in general, we examined diagnoses of merely distressed behaviors. In Experiment 1, traumatic event context increased disorder diagnoses for the non-disordered distressed behaviors. However, we didn’t find this to be the case in Experiment 2; thus, it remains somewhat unclear whether traumatic event context affects clinical judgments about non-disordered behaviors, at least all of the time. Future research will need to further explore the conditions under which traumatic event context matters for clinical evaluations of more everyday behaviors.

Our fourth question was whether we would replicate psychological abnormality and difficulty-of-understanding judgments that supported the proportionate-response effect, as in past work. In Experiment 1, we cleanly replicated these effects.

Question 5 asked whether there were any effects of expertise. There were none in Experiment 1. In Experiment 2, we found that clinical trainees had better sensitivity in the recognition task and faster reaction times overall compared to clinicians. On further exploration, we found that sensitivity was negatively correlated with years of clinical experience and reaction time was positively correlated with years of clinical experience; that is, more expertise made
clinicians less able to distinguish false from true patient descriptions and slower to make these decisions.

Sixth, we wanted to investigate why reactions did not appear to influence any judgments in Experiment 1; namely, we tested whether participants were especially susceptible to forming distorted memories for reactions. Our data were consistent with this possibility. We found that participants made more false alarms for reactions and had poorer sensitivity for reactions compared to either events or behaviors, whereas hit rates for all three were comparable.

Seventh, we also wanted to know whether people’s memory of the reaction information was distorted, such that the strength of the reaction information was recollected as better matching the strength of the event or behaviors than it actually did in the vignette. Consistent with this possibility, we found evidence that participants made more false alarms for false strength reaction items in vignettes where the reactions were disproportionate to the event or the behaviors than they did when the reactions were proportionate to the event or behaviors in strength.

**Lack of Influence of Reactions**

The data collected to answer Questions 6 and 7 indicated that participants were susceptible to remembering the reactions as consistent with the events and behaviors, even if they were not consistent with one another in the original vignette. This finding supported the possibility that reactions did not influence clinical judgments, including diagnoses, at least in part because clinicians did not necessarily remember them accurately, distorting them to be aligned with the events or behaviors.

The alternative hypothesis that participants were not influenced by reactions because they remembered them but simply did not find them to be relevant for any of the questions we asked
was not fully supported by the current data. If this had been the case, then participants should have had equivalent memory for reactions as for events or behaviors, but they did not.

**Clinical Conceptualizations of PTSD**

Given all of the information gathered about PTSD in this study, we were in a good position to speculate about the likely structure of clinician’s concepts of PTSD. We speculate that the event and behaviors are very central features in the PTSD concept. Both the event and behaviors were so important that they affected people’s memory for the reactions, such that reactions were remembered as being more closely in alignment with the event and behaviors than they actually were. As we have discussed, PTSD is a unique entity amongst mental disorders in the *DSM-5* (APA, 2013), in that context is part of the PTSD diagnostic criteria in the form of a traumatic event. Not only were clinicians in our studies influenced by the event when identifying cases of PTSD behaviors as PTSD, they sometimes even identified PTSD for non-PTSD behavioral symptoms presented in traumatic event context. Of course, as expected, the PTSD behavioral symptoms also played a very important role, guiding people to make far more PTSD diagnoses for vignettes with PTSD behaviors than for those with MDD or distressed behaviors (please see Table B3). Also, people made the fewest false alarms for behaviors, indicating their importance to clinicians when learning about the hypothetical people described in the vignettes.

The reaction appears likely to be a very peripheral feature in clinicians’ PTSD concepts, if it is included at all. Reactions did not seem to be an important consideration for participants, with people making the most false alarms for reactions and misremembering them such that reactions were matched to the strength of the event and/or behaviors. It is interesting to speculate whether the perceived lack of salience of reactions might not only interfere with its ability to influence judgments, but also make it difficult for clinicians to fully internalize research
indicating the particular importance of reactions in precipitating PTSD behavioral symptoms (e.g., Boals & Schuettler, 2009; Bodkin, Pope, Detke, & Hudson 2007; Gold, Marx, Soler-Baillo, & Sloan 2005; Rubin & Feeling, 2013). The latter possibility remains to be explored.

**Clinical Utility of the DSM-5**

Interestingly, people diagnosed PTSD according to the DSM-5 (APA, 2013) criteria regardless of their reported experience with that version of the manual. About half of all Experiment 1 participants reported no or minimal DSM-5 experience. How then, did their diagnoses align with it anyway? One possibility is that clinicians did not base their diagnoses on the DSM-5 (APA, 2013), but conversely, the DSM-5’s (APA, 2013) diagnostic criteria were based on the preferences of those working in clinical psychology. Indeed, the DSM-5 (APA, 2013) claims that clinical utility (i.e., the extent to which the DSM helps clinical professionals understand, diagnose, and communicate about disorders, make treatment decisions, and predict future needs; First et al., 2004) is one of its primary goals. To reach this goal, the DSM-5 (APA, 2013) appears to have shaped its diagnostic criteria to fit clinical intuitions about disorders.

As we discussed in the opening of this paper, there was much controversy surrounding the description of PTSD in the DSM-5 (APA, 2013; Hathaway et al., 2010). For example, some research demonstrated the usefulness of the A2 criterion in detecting PTSD symptomology (e.g., Brewin et al., 2000); other research demonstrated that the A2 criterion was superfluous to PTSD symptomology (e.g., Karam et al., 2010). Given multiple options for the diagnostic criteria for PTSD in the DSM-5 (APA, 2013), supported by different lines of research, the DSM-5 task force (APA, 2013) chose to implement the version of PTSD that includes traumatic event context only (the exact justifications for their final decisions apparently remain relatively mysterious to those outside the task force; Whooley, 2014). That is, given many options, the DSM-5 task force
(APA, 2013) chose to define PTSD using the kind of context that made clinicians in this study, paradoxically, view PTSD behaviors as the least psychologically abnormal and the least difficult to understand.

Ultimately, it is mental health clinicians who serve on the DSM task forces and decide on the criteria that appear in the *DSM*. Our data are consistent with the broad speculation that the research supporting the A2 criterion was ultimately unsuccessful in shaping PTSD in the *DSM-5* (APA, 2013) in part because clinicians do not intuitively feel that intense reaction is a part of PTSD. That is, cognitive difficulty in understanding how an intense reaction could cause PTSD could have helped prevent such research from being perceived as compelling enough to influence the manual. Future *DSM* task forces might benefit from some degree of reflection upon whether cognitive difficulty in accepting the particular implications of new research has a tendency to shape new diagnostic criteria in the manual.

**What do Mental Disorder Experts Think Constitutes a Mental Disorder?**

Is mental disorder about psychological abnormality? Our clinicians did not seem to think so. That is, the severely negative behaviors that clinicians found to be the least psychologically abnormal (disordered behavioral symptoms presented in traumatic event context) were diagnosed more often with PTSD and no less often with depression than the disordered behaviors they found to be comparatively more psychologically abnormal. That is, it did not matter to clinicians whether they thought disordered behaviors were a psychologically normal response to a recent life event. When asked, they labeled even responses that they personally judged to be relatively psychologically normal as disordered.

Under one view, disordered behavioral symptoms that are an appropriate response to environmental circumstances should not be categorized as disordered (Horwitz & Wakefield,
Rather, in this view, behavioral symptoms are often an appropriate and healthy response to a negative environmental change (e.g., a recent major loss), which elicits a period of personal mental distress (e.g., deep sadness while coping with the loss), yet incurs a benefit in the long run (e.g., effective navigation of changed life circumstances). In this view, interrupting this process with inappropriate treatment is harmful because it interferes with needed adaptation. On the other hand, one argue that disordered behavioral symptoms involve the inappropriate activation of such an inner response to adversity, which in turn causes needless personal distress, which over time will not remit on its own and therefore requires clinical treatment (Horwitz & Wakefield, 2007).

Although Horwitz & Wakefield (2007) make a clear case for distinguishing normal distress from disorder, the DSM is not structured to specify how clinicians should distinguish the two. Ever since the DSM-III (APA, 1980), diagnosis has been limited to symptom counting, excluding any type of causal information (with the rare exception of PTSD, wherein the cause is included as a symptom itself). As we have found here, this diagnostic system was effective in eliminating causal information not only from the diagnostic criteria themselves, but also from the realm of clinical consideration. That is, the way the DSM is now structured encourages clinicians to diagnose all behavioral symptoms that fulfill given symptom criteria, whether they find those symptoms to be an appropriate coping response.

According to Horwitz & Wakefield (2007), the DSM improperly and harmfully causes over-diagnosis of disorders in an attempt to bring as many distressed individuals as possible within the reach of clinical services. In their view, the intuitive judgments of clinicians, that some behavioral symptoms represent a coping response and therefore are non-disordered, should serve to drive down the number of people whom might receive a diagnosis. The symptom
counting system of diagnosis, introduced in the *DSM-III* (APA, 1980) seems to have found a way around this possibility. That is, the modern *DSM* structures diagnosis such that it circumvents clinicians’ intuitions, making it unnecessary to think about the relationship between life events and behaviors in order to enter a diagnosis decision. Specifically, it does this by hinging diagnoses on the identification of a set of behavioral symptoms, an assessment which does not explicitly specify whether, and how, causal context ought to be considered (with the rare exceptions explored in the current research).

**The Influence of Traumatic Event Context: Increasing PTSD Diagnoses, Decreasing Psychological Abnormality Judgments**

What is mental disorder to clinicians, then, if not psychological abnormality? Our diagnosis findings do not align with the classic notion in abnormal psychology that abnormal behaviors merit a diagnosis but normal behaviors do not (Bennett, 2003; Bennett, 2011; Horwitz & Wakefield, 2007). Under the assumption that clinicians would tend to hold this perspective, we would have expected traumatic event context to influence both PTSD diagnoses and psychological abnormality judgments in the same direction, by increasing both or decreasing both. Yet it is plausible that a number of different factors could influence diagnostic judgments at the same time, of which perceived abnormality is only one. That is, the diagnosis judgments elicited in this study might be the net effect of this hypothetical set of factors. According to the *DSM-5* (APA, 2013), the perception that behaviors are psychologically abnormal (an assessment which, too, may be informed by a number of things) is one factor that points to the need for a diagnosis. But as these data have shown so far, psychological abnormality judgments can run counter to diagnosis judgments.
What might some of these additional factors be? It has been argued in philosophical discussions of mental health diagnosis that behaviors that cause an individual distress and/or disability are behaviors that merit treatment (Bennett, 2011), and are some of the most important reasons why a person might need a diagnosis (Kring, Johnson, Davison, & Neale, 2010). Distress is characterized by suffering experienced by an individual. It can manifest as a negative internal state or an emotionally painful experience, and can include guilt, remorse, and anxiety. On the other hand, disability is characterized by impairment in any major area of functioning in life. This impairment can range from social rejection, which limits interpersonal interactions, to impairment in physical mobility, which limits daily activities, fulfillment of key roles, and an individual’s ability to function independently (Kring et al., 2010). Finally, treatment need refers to behaviors that merit alleviation, or repair, through intervention and care (Bennett, 2003). It would be reasonable for clinicians to consider such factors alongside that of pure psychological abnormality in making diagnoses. In an exploratory experiment (please see Appendix E), we took an initial step toward addressing this possibility directly.

Limitations

First, this study sample was limited to clinicians and clinical trainees who self-selected to participate. Thus, all study participants had sufficient interest in the topics broached in the recruitment materials (e.g., research on clinical decision making), and / or were interested in the token of thanks offered for participation (a $25 gift card to an online retailer for clinicians, a raffle entry for a $25 gift card to an online retailer for clinical trainees). Although all participants had some interest in the study before participating, this is the case with much experimental research, wherein study participation is always voluntary.
Second, people in this study judged hypothetical others described in vignettes, albeit as realistically as possible, rather than learning about actual people’s recent past and subsequent behaviors in person, as in clinical practice. However, it was necessary to conduct this research with vignettes so that we could carefully control the events, reactions, and behaviors described. Whether these findings would be replicated in more ecologically valid situations should be tested in future work.

Third, although we found clear effects of context on judgments of psychological abnormality, it is likely that participants had different ideas about the exact meaning of the term. The construct of psychological abnormality has no singular, explicit definition, but rather denotes the general idea that psychologically abnormal behaviors are those that merit assessment and treatment (Bennett, 2011). We measured participants’ judgments of psychological abnormality not under the assumption that there is one correct and widely used definition, but because of the common framework it denotes in clinical psychology, within which many clinicians cultivate their own meaning through experience. (Please see Appendix E for a more extensive look at this issue).

Fourth, our data do not indicate whether the effects we found with recognition items occurred during encoding, retrieval, or storage. That is, this experiment does not provide detailed information about which aspect of memory is responsible for the effects described here. However, it does tell us that some part of the process of encoding, storing, and/or retrieving memories was distorted for reactions selectively. This finding constitutes an important first step toward understanding how one type of proportionate, causal context (i.e., the event) can influence clinical judgments while another (i.e., the reaction) can fail to do so.

Fifth, it is possible that clinicians did not remember the names of people in the vignettes,
which was the primary cue. However, given the many ceiling effects we found, this possibility seems unlikely. That said, in future work, it would be useful to provide participants with a brief reminder of the vignette (in addition to the name), which does not indicate the correct answer to any of the recognition items but still brings the vignette to mind.

Finally, collecting reaction time data online involved uncontrolled sources of error and variance that could have obscured underlying reaction time effects, such as differences in computer processing speed and variance in internet connection, though these potential issues were addressed to the extent possible by trimming outliers. Given the generally high level of difficulty in recruiting clinicians, this method was a good way of attempting to collect reaction time data for a first look, and certainly was suitable for collecting the accuracy data, which were by far most critical to testing our hypotheses. Future work could investigate clinicians’ reaction times for recognition questions in person.

Conclusion

This work allows us to achieve a deeper understanding of how clinicians and clinical trainees assess mental disorders in context, in an era where de-contextualization (Whooley, 2014) and over-diagnosis (Horwitz & Wakefield, 2007) have become the norm. More specifically, it allowed us to look at the particular case wherein context is a diagnostic criterion, rather than merely background information about a client that professionals might consider. Our work suggests that including context in the diagnostic criteria, as is the case for PTSD, leads to different clinical and diagnostic judgments than when it is not. Many observers of the DSM system have argued that de-contextualization leads to inappropriate over-diagnosis (e.g., a person grieving the death of a loved one can still legitimately receive a MDD diagnosis in DSM-5; APA, 2013, and it has been argued that this is wholly unreasonable; Horwitz & Wakefield, 2007). The
current work suggests that one form of formal contextualization, the inclusion of context as a diagnostic criterion in the DSM, does not decrease diagnoses in such context, as some assert would be appropriate, but rather increases diagnoses. That is, even though some might think that the inclusion of context in PTSD diagnostic criteria addresses the de-contextualization problem of over-diagnosis, our experiments suggest that to the contrary, adding context to the diagnostic criteria for a disorder might in fact do the opposite.
References


Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201-208).

Cambridge, MA: Cambridge University Press.


### Experiment 1 Partial Sample Vignette – “Recent Past”

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Reaction Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traumatic Event</td>
<td><strong>Intense Reaction (in italics)</strong>&lt;br&gt;Lucy teaches at an elementary school where there was a school shooting several months ago. She was with her class of children in the playground when they were pinned down by sniper fire. <em>Immobilized and vulnerable she became completely paralyzed with fear. She felt like they were all going to die.</em>&lt;br&gt;Lucy teaches at an elementary school where there was a school shooting several months ago. She was with her class of children in the playground when they were pinned down by sniper fire. <em>She was caught off guard and confused about what was going to happen. She tried to be confident for the children.</em></td>
</tr>
<tr>
<td>Everyday Event</td>
<td><strong>Mild Reaction (in italics)</strong>&lt;br&gt;Lucy teaches at an elementary school where there was a lockdown drill several months ago. She hadn’t known about the drill in advance because she had been out sick and missed the memo. <em>Immobilized and vulnerable she became completely paralyzed with fear. She felt like they were all going to die.</em>&lt;br&gt;Lucy teaches at an elementary school where there was a lockdown drill several months ago. She hadn’t known about the drill in advance because she had been out sick and missed the memo. <em>She was caught off guard and confused about what was going to happen. She tried to be confident for the children.</em></td>
</tr>
</tbody>
</table>

*Note: This is a partial vignette for illustrative purposes. The actual vignettes described 4-5 times.*
more information (see Appendix A for complete text Reaction text, italicized here to differentiate it from event text, was not italicized in the main study.
Table B2

*Experiment 1 Dependent Variables*

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Question Text</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open-Ended</td>
<td>Given [X]’s subsequent behaviors, what mental disorder(s) is [X] likely to have, if any?</td>
<td>Please respond to this question by typing the name of a mental disorder into Diagnosis Box 1 below.¹</td>
</tr>
<tr>
<td>Diagnoses</td>
<td>How psychologically normal, or psychologically abnormal, are [X]’s subsequent behaviors?</td>
<td>Please choose your response on a scale of 1-9, where 1=very psychologically abnormal and 9=very psychologically normal.</td>
</tr>
<tr>
<td>Psychological Abnormality</td>
<td>How easy or difficult is it to understand [X]’s subsequent behaviors?</td>
<td>Please choose your response on a scale of 1-9, where 1=very easy to understand and 9=very difficult to understand.</td>
</tr>
</tbody>
</table>

¹Open-ended diagnoses instructions continued: If you believe no diagnosis is warranted, please type “none” into Diagnosis Box 1, and leave the other boxes blank. If there are any other diagnoses you think are likely, please type them into the subsequent Diagnosis Boxes 2 and 3 (one diagnosis per box). Then, please rate the likelihood of each diagnosis on a scale of 1-9, where 1=very unlikely and 9=very likely, by selecting a rating on the scale below each diagnosis box. If you believe no diagnosis is warranted, please also rate the likelihood of this answer on the scale of 1-9, where 1 = very likely and 9 = very unlikely.
<table>
<thead>
<tr>
<th>Open-Ended Diagnosis Responses</th>
<th>PTSD</th>
<th>MDD</th>
<th>Distressed</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTSD</td>
<td>192</td>
<td>73</td>
<td>11</td>
</tr>
<tr>
<td>Depression (including MDD)</td>
<td>58</td>
<td>249</td>
<td>9</td>
</tr>
<tr>
<td>&quot;None&quot;</td>
<td>9</td>
<td>2</td>
<td>233</td>
</tr>
<tr>
<td>Adjustment Disorder</td>
<td>28</td>
<td>32</td>
<td>18</td>
</tr>
<tr>
<td>Anxiety Disorder</td>
<td>26</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Acute Stress Disorder</td>
<td>20</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>Generalized Anxiety Disorder</td>
<td>18</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Bipolar Disorder</td>
<td>9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Panic Disorder</td>
<td>7</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Dysthymia</td>
<td>1</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personality Disorder</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Binge Eating Disorder</td>
<td></td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Specific Phobia</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eating Disorder</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borderline Personality Disorder</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Mood Disorder</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Obsessive Compulsive Disorder</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disorder</td>
<td>DSM-IV-TR</td>
<td>DSM-5</td>
<td>DSM-5TR</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>-----------</td>
<td>-------</td>
<td>---------</td>
</tr>
<tr>
<td>Social Anxiety Disorder</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Unspecified Trauma and Stressor Related Disorder</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Insomnia Disorder</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Depersonalization Disorder</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissociative Disorder</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizoaffective Disorder</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non DSM-IV-TR or DSM-5 Disorder Responses</td>
<td>8</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total Responses</strong></td>
<td><strong>405</strong></td>
<td><strong>407</strong></td>
<td><strong>300</strong></td>
</tr>
</tbody>
</table>
Table B4

Experiment 1 Mean Ratings for Psychological Abnormality Judgments and Difficulty Understanding Judgments by Context and Behavior Type, with Statistical Significance of Event Type Main Effects from 2x2 ANOVAs

<table>
<thead>
<tr>
<th>Behavior and Judgment Type</th>
<th>Traumatic Event Mean (SD)</th>
<th>Everyday Event Mean (SD)</th>
<th>Event Type p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychological Abnormality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTSD</td>
<td>5.10 (2.12)</td>
<td>6.60 (1.60)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>MDD</td>
<td>4.76 (2.10)</td>
<td>6.75 (1.45)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Distressed</td>
<td>2.71 (1.39)</td>
<td>2.20 (1.17)</td>
<td>= .042</td>
</tr>
<tr>
<td>Difficulty Understanding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTSD</td>
<td>3.03 (1.82)</td>
<td>2.26 (1.27)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>MDD</td>
<td>2.53 (1.26)</td>
<td>5.13 (2.14)</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>Distressed</td>
<td>3.03 (1.82)</td>
<td>2.26 (1.27)</td>
<td>= .007</td>
</tr>
</tbody>
</table>

Note. Psychological Abnormality judgments were made on a scale of 1-9, where 1 = very psychologically normal, and 9 = very psychologically abnormal. Difficulty Understanding judgments were made on a scale of 1-9, where, 1 = very easy to understand, and 9 = very difficult to understand.
### Experiment 2 Event Recognition Items

<table>
<thead>
<tr>
<th>Event Type and Recognition Item</th>
<th>Vignette Name</th>
<th>Martin</th>
<th>Lucy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traumatic Event</td>
<td>Cathy</td>
<td>Martin got hit by a tractor-trailer</td>
<td>Lucy wasn't able to protect her students</td>
</tr>
<tr>
<td>Original Correct 1</td>
<td>There was an explosion outside town</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original Correct 2</td>
<td>Cathy was at the hospital when it happened</td>
<td>The passenger's side was impacted</td>
<td>One student was killed</td>
</tr>
<tr>
<td>Original Correct 3</td>
<td>Distraught families came to the hospital</td>
<td>His daughter was badly hurt</td>
<td>Police found the sniper at his home</td>
</tr>
<tr>
<td>Original Correct 4</td>
<td>Cathy went without sleep for many hours</td>
<td>Martin was not harmed in the accident</td>
<td>The sniper committed suicide</td>
</tr>
<tr>
<td>False Content 1</td>
<td>Cathy's friend was among the wounded</td>
<td>Fire trucks arrived at the scene</td>
<td>The sniper called in a bomb threat</td>
</tr>
<tr>
<td>False Content 2</td>
<td>The hospital briefly lost power</td>
<td>Martin's son had a broken arm</td>
<td>Police took 30 minutes to find the sniper</td>
</tr>
<tr>
<td>False Strength 1</td>
<td>Injured people trickled in slowly</td>
<td>The tractor-trailer merely tapped</td>
<td>Only two students were injured</td>
</tr>
</tbody>
</table>
Martin's car

False Strength 2

Everyone received treatment eventually His daughter healed rapidly Just after the gunfire ceased, people got up

Everyday Event

Original Correct 1

Cathy listened to a talk about military medics Martin was at a 4-way stop Lucy was unaware of the scheduled drill

Original Correct 2

She saw pictures of injuries from a bomb The accident made a loud sound Everyone got under their desks

Original Correct 3

The officers talked about time and equipment Martin saw one of the cars run a stop sign Lucy tried to calm her students

Original Correct 4

She heard medics cannot help everyone who is dying No one was injured in the accident An officer announced it was over

False Content 1

Cathy watched a mock emergency procedure One of the drivers called the police One student started screaming

False Content 2

She heard that many become traumatized Martin got out of his car to help Lucy locked her classroom door

False Strength 1

Cathy learned almost Someone from each All of the children
no one survives these bombs

False Strength 2

The photos were the most graphic she has seen

Martin thought the drivers would fight

Many cops swarmed into the class

car was injured
were utterly terrified
### Table B6

**Experiment 2 Reaction Recognition Items**

<table>
<thead>
<tr>
<th>Vignette Name</th>
<th>Reaction Type and Recognition Item</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cathy</td>
</tr>
<tr>
<td></td>
<td>Martin</td>
</tr>
<tr>
<td></td>
<td>Lucy</td>
</tr>
<tr>
<td>Intense Reaction</td>
<td></td>
</tr>
<tr>
<td><strong>Original Correct 1</strong></td>
<td>Cathy experienced complete horror</td>
</tr>
<tr>
<td></td>
<td>Martin felt extreme stress in the moment</td>
</tr>
<tr>
<td></td>
<td>Lucy was extremely afraid</td>
</tr>
<tr>
<td><strong>Original Correct 2</strong></td>
<td>She went into a state of disbelief</td>
</tr>
<tr>
<td></td>
<td>His heart rate went up rapidly</td>
</tr>
<tr>
<td></td>
<td>She prayed for it to end</td>
</tr>
<tr>
<td><strong>Original Correct 3</strong></td>
<td>Cathy felt obligated to stay until it was over</td>
</tr>
<tr>
<td></td>
<td>This was Martin's worst experience by far</td>
</tr>
<tr>
<td></td>
<td>Lucy was not relieved</td>
</tr>
<tr>
<td><strong>Original Correct 4</strong></td>
<td>She started shaking and couldn't stop</td>
</tr>
<tr>
<td></td>
<td>Martin couldn't move from fear</td>
</tr>
<tr>
<td></td>
<td>She realized schools can be attacked</td>
</tr>
<tr>
<td><strong>False Content 1</strong></td>
<td>Cathy started breathing very quickly</td>
</tr>
<tr>
<td></td>
<td>Martin closed his eyes tightly</td>
</tr>
<tr>
<td></td>
<td>She was shaking violently</td>
</tr>
<tr>
<td><strong>False Content 2</strong></td>
<td>She nearly fainted more than once</td>
</tr>
<tr>
<td></td>
<td>He kept saying &quot;no&quot; over and over</td>
</tr>
<tr>
<td></td>
<td>Lucy noticed she was screaming</td>
</tr>
<tr>
<td><strong>False Strength 1</strong></td>
<td>Cathy noticed she was holding it together</td>
</tr>
<tr>
<td></td>
<td>Internally, Martin felt a little surprised</td>
</tr>
<tr>
<td></td>
<td>Lucy hoped everything would turn out okay</td>
</tr>
</tbody>
</table>
She took a break to center herself. Martin felt slightly uneasy. She felt she handled it the best she could.

Mild Reaction

Cathy's experience was jarring. Martin was rather surprised. Lucy found it hard to concentrate.

She frowned. He held his breath for a moment. She knew she did the best she could.

Cathy felt sympathy for others in pain. Martin knew it could have been prevented. Lucy felt relief once it ended.

She became fatigued by this experience. He hadn't expected this to happen. She realized schools are vulnerable.

Cathy realized she was confused. Martin cursed under his breath. She was angry that this was happening.

She thought about going home. He thought about all the bad drivers in the world. Lucy was grateful she had emergency training.

She quickly began to panic. He thought someone had died in the accident. She nearly had a panic attack.

Cathy was in tears thinking about the event. Martin remained on edge for days. This gave Lucy a migraine.
### Table B7

**Experiment 2 Behaviors Recognition Items**

<table>
<thead>
<tr>
<th>Behavior Type and Recognition Item</th>
<th>Vignette Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTSD Behaviors</td>
<td></td>
</tr>
<tr>
<td>Original Correct 1</td>
<td></td>
</tr>
<tr>
<td>Cathy could often be found crying</td>
<td>Cathy</td>
</tr>
<tr>
<td>Martin stopped caring about his life</td>
<td>Martin</td>
</tr>
<tr>
<td>Lucy had a lasting sense of hopelessness</td>
<td>Lucy</td>
</tr>
<tr>
<td>Original Correct 2</td>
<td></td>
</tr>
<tr>
<td>She had a hard time sleeping through the night</td>
<td>Cathy</td>
</tr>
<tr>
<td>He began doing poorly at work</td>
<td>Martin</td>
</tr>
<tr>
<td>She felt exhausted all of the time</td>
<td>Lucy</td>
</tr>
<tr>
<td>Original Correct 3</td>
<td></td>
</tr>
<tr>
<td>Generally she felt empty inside</td>
<td>Cathy</td>
</tr>
<tr>
<td>Martin felt badly about the incident</td>
<td>Martin</td>
</tr>
<tr>
<td>Lucy began doing poorly at work</td>
<td>Lucy</td>
</tr>
<tr>
<td>Original Correct 4</td>
<td></td>
</tr>
<tr>
<td>Cathy noticed herself speaking slowly sometimes</td>
<td>Cathy</td>
</tr>
<tr>
<td>He stopped eating as much as he used to effectively</td>
<td>Martin</td>
</tr>
<tr>
<td>She could not concentrate</td>
<td>Lucy</td>
</tr>
<tr>
<td>False Content 1</td>
<td></td>
</tr>
<tr>
<td>Cathy picked at her food at meals</td>
<td>Cathy</td>
</tr>
<tr>
<td>Martin kept mumbling to himself</td>
<td>Martin</td>
</tr>
<tr>
<td>Lucy felt angry with herself</td>
<td>Lucy</td>
</tr>
<tr>
<td>False Content 2</td>
<td></td>
</tr>
<tr>
<td>She would stare at an empty wall for hours</td>
<td>Cathy</td>
</tr>
<tr>
<td>He lay in bed for hours on end</td>
<td>Martin</td>
</tr>
<tr>
<td>She avoided seeing family members</td>
<td>Lucy</td>
</tr>
<tr>
<td>False Strength 1</td>
<td></td>
</tr>
<tr>
<td>Cathy had about as much energy as</td>
<td>Cathy</td>
</tr>
<tr>
<td>Martin's mind rarely wandered</td>
<td>Martin</td>
</tr>
<tr>
<td>She felt sad only on rare occasions</td>
<td>Lucy</td>
</tr>
</tbody>
</table>
She only rarely felt empty or sad. His boss thought he was doing all right most of the time. Lucy enjoyed herself most of the time.

**MDD Behaviors**

**Original Correct 1**
- Cathy kept reimagining that day at home. Lucy felt a lasting sense of shame.

**Original Correct 2**
- Her work began to suffer. He had nightmares of the incident. She couldn't escape thoughts of the event.

**Original Correct 3**
- Cathy's ability to concentrate was diminished. Martin had spurts of rage. Lucy had a short temper.

**Original Correct 4**
- She avoided talking about what happened. He was unable to control himself. She didn’t want to talk about it with anyone.

**False Content 1**
- Cathy had a hard time eating anything. Martin felt he was a terrible person. Lucy hated the way she looked.

**False Content 2**
- She sometimes wished she was dead. He scratched his skin convulsively. She started drinking a lot of alcohol.

**False Strength 1**
- She usually participated in her hobbies. Martin maintained his relationships. Just once, Lucy felt bad about what happened.

**False Strength 2**
- Cathy wasn't too worried about the sleeping. He rarely had trouble. She improved somewhat over the
Cathy sometimes felt a little on edge. Martin had occasional anxiety. Lucy sometimes felt uneasy.

She thought sadly about world violence. He was very attentive to the road. One night she felt wide awake.

Cathy felt good about her life. Martin felt generally okay. She talked with friends and co-workers.

She was calm during her usual routines. He wanted to keep his kids safe. Sometimes Lucy felt a little anxious.

False Content 1
She sometimes felt tired on the weekend. Martin considered getting a new car. Lucy took walks to clear her mind.

False Content 2
Cathy kept in touch with supportive friends. He often thought about the future. She stayed in close touch with her mother.

False Strength 1
Cathy’s work routine made her fearful crossing intersections. Months later, she felt quite paranoid.

False Strength 2
She often felt disaster was pending. Martin was often on the edge of panic. She sometimes lashed out at friends.
Figure B1

Experiment 1 Clinicians’ and Clinical Trainees’ Reported DSM-5 Experience

![Bar chart showing the reported DSM-5 experience of clinicians and clinical trainees.](chart.png)
Figure B2. Each column displays the number of diagnoses of interest for 136 vignettes. Sixty-eight participants were in the traumatic event condition, and 68 in the everyday event condition. Each participant provided diagnoses for two vignettes describing PTSD behaviors, two describing MDD behaviors, and two describing distressed behaviors. Thus, a total of 136 vignettes per column received diagnoses.
Figure B3

Experiment 2 Open-Ended Diagnosis Encoding Task Results

![Graph showing the number of diagnoses for PTSD, Depression, and "None" behaviors differentiated by traumatic event and everyday event.](image-url)
Figure B4

Experiment 2 Manipulation Check Severity Ratings

![Graph showing severity ratings for severe and mild vignette elements with correct original, false content, and false strength conditions.]
Figure B5

Experiment 2 Proportion of False Alarms for Events, Reactions, and Behaviors
Figure B6

Experiment 2 Proportion of Hits for Events, Reactions, and Behaviors
Figure B7

Experiment 2 Proportion of False Alarms for False Strength Reaction Items in Proportionate and Disproportionate Reaction and Behaviors Conditions
Appendix A

Full Vignettes

In this Appendix, the full vignettes used in the work with college students and the work with clinicians and clinical trainees are listed.\(^5\) The vignettes describe six hypothetical people in total: Dara, Cathy, Martin, Justin, Lucy, and Diana. Within each hypothetical person’s section, the “Recent Past” information is shown first, followed by the “Subsequent Behaviors” information, as in the experimental procedures. For each hypothetical person, there are four “Recent Past” paragraphs, each presenting one of the four combinations of event type and reaction type: traumatic event / intense reaction, traumatic event / mild reaction, everyday event / intense reaction, everyday event / mild reaction, in that order. Within each “Recent Past” paragraph, introductory background material is underlined, event information is in regular font, and reaction information is in italics.

The “Recent Past” paragraphs are followed by the “Subsequent Behaviors” information, which consists of three paragraphs describing that hypothetical person’s PTSD behavioral vignette, MDD behavioral vignette, and distressed behavioral vignette, in that order. Word counts are provided at the end of each vignette section.

\(^5\) The vignettes employed in the college student paper and the vignettes employed in the clinician paper differ very slightly. After completing the work on college students but before beginning the work performed with clinicians and clinical trainees, a practicing clinical psychologist was asked to review the vignettes for accuracy and clarity. Several details were altered slightly as a result. All edits, detailed in this Appendix, were made to decrease the likelihood that one symptom pointed towards a different diagnosis than intended. These changes were only made to MDD and distressed behavior vignettes. The materials in this Appendix reflect these changes.
Dara’s “Recent Past” Vignettes

**Dara’s traumatic event, intense reaction “recent past” vignette.** Dara was born in a rural area of Darfur and grew up on her family’s farm. When the Janjaweed came to power under al-Bashir when Dara was a teenager, she was separated from her family. *This deeply shocked Dara.* As the reality of what happened sank in, she felt paralyzed. She tried to understand why this had happened and felt completely alone. One of her brothers was executed. *Dara was devastated.* She thought that terrible people fill the world, and cannot be stopped. She reunited with her mother, but only briefly, as her mother died of starvation several hours later. *This broke Dara’s spirit.* She felt weak with terror and defeated. *The world she knew was falling apart.* She felt defenseless and unprotected. She was then forced to marry, and shortly after, relocated to a refugee camp. There, she had very little food to eat and saw many people starve to death. Others were killed outright. *Dara felt completely helpless about what she had seen.* She thought she could never do enough good in her life to make up for this suffering. In 2010, Dara left the refugee camp on her own and was able to escape the country. *Even so, Dara felt overwhelmingly helpless.* *She had the impression that her life was very small compared to all of those lives that had been lost.* She also felt powerless and weak. She fled to Egypt, later moved to the U.S., and has never returned to Darfur. (Event: 112 words, Reaction: 128 words)

**Dara’s traumatic event, mild reaction “recent past” vignette.** Dara was born in a rural area of Darfur and grew up on her family’s farm. When the Janjaweed came to power under al-Bashir when Dara was a teenager, she was separated from her family. *As the reality of what happened sank in, she felt strange.* She tried to understand why this had happened, but it didn’t make much sense. *This all was quite unexpected, and felt a bit surreal.* One of her brothers was executed. *This made Dara wish she had someone to talk about it.* She reunited with her
mother, but only briefly, as her mother died of starvation several hours later. Dara was 
dismayed. She was reminded that the world can be a dark place. She was then forced to marry, 
and shortly after, relocated to a refugee camp. There, she had very little food to eat and saw 
many people starve to death. Others were killed outright. Dara thought what she saw was 
unfortunate, but knew she could do nothing to change it. She could not help but wish that one of 
the people in power would do something to help. In 2010, Dara left the refugee camp on her own 
and was able to escape the country. Dara felt moved. She realized that people can be cruel, and 
the world is not always predictable. She thought that some people are just impossible to reason 
with. She fled to Egypt, later moved to the U.S., and has never returned to Darfur. (Event: 112 
words, Reaction: 128 words)

Dara’s everyday event, intense reaction “recent past” vignette. Dara read a newspaper 
article about the Killing Fields. The Killing Fields is a name given to the genocide that took 
place in Cambodia in the 1970s. This deeply shocked Dara. As the reality of what happened sank 
in, she felt paralyzed. She tried to understand why this had happened and felt completely alone. 
She read that the Khmer Rouge regime, lead by Pol Pot, held power for nearly four years, during 
which about 2 million Cambodians died from starvation, torture, or execution. Dara was 
devastated. She thought that terrible people fill the world, and cannot be stopped. People were 
killed because of their religion, culture, social or economic group; some survivors lost their 
entire families. This broke Dara’s spirit. She felt weak with terror and defeated. The world she 
knew was falling apart. She felt defenseless and unprotected. The article described the history of 
the Killing Fields and the events in Cambodia during this time period, and Dara saw many 
photographs of what happened. Dara felt completely helpless about what she had seen. She 
thought she could never do enough good in her life to make up for this suffering. She read that
The Vietnamese invaded in 1979, and many people escaped the country. Even so, Dara felt overwhelmingly helpless. She had the impression that her life was very small compared to all of those lives that had been lost. She also felt powerless and weak. Neither Dara nor anyone she knows was personally affected. It is not the case that Dara reacted this way because she was reminded of something in her past; she was truly this upset by the article alone. (Event: 123 words, Reaction: 156 words)

Dara’s everyday event, mild reaction “recent past” vignette. Dara read a newspaper article about the Killing Fields. The Killing Fields is a name given to the genocide that took place in Cambodia in the 1970s. As the reality of what happened sank in, she felt strange. She tried to understand why this had happened, but it didn’t make much sense. This all was quite unexpected, and felt a bit surreal. She read that the Khmer Rouge regime, lead by Pol Pot, held power for nearly four years, during which about 2 million Cambodians died from starvation, torture, or execution. This made Dara wish she had someone to talk to about it. People were killed because of their religion, culture, social or economic group; some survivors lost their entire families. Dara was dismayed. She was reminded that the world can be a dark place. The article described the history of the Killing Fields and the events in Cambodia during this time period, and Dara saw many photographs of what happened. Dara thought what she read was unfortunate, but knew she could do nothing to change it. She could not help but wish that one of the people in power had done something to help. She read that the Vietnamese invaded in 1979, and many people escaped the country. Dara felt moved. She realized that people can be cruel, and the world is not always predictable. She thought that some people are just impossible to reason with. Neither Dara nor anyone she knows was personally affected. (Event: 123 words, Reaction: 128 words)
Dara's Behavioral Vignettes: “Subsequent Behaviors”

Dara’s PTSD behavioral vignette. For several months later, Dara screamed and shook whenever she heard a loud noise or was surprised. Dara kept seeing those bad things over and over in her mind. She also felt separated from the people around her, as if they were living in another world – or she was. She avoided people whom she thought might remind her of her past. Now, she still feels detached from others, and is tormented by those bad images in her mind. Dara continues to scream and shake when surprised. (86 words)

Dara’s MDD behavioral vignette. For several months later, Dara had a hard time making simple decisions, such as figuring out what to eat. She felt intensely sad all of the time. She often thought about death’s inevitability. Dara felt awful inside and just wanted to sleep. She believed that her life was worthless. Most of the time she felt exhausted, even though she slept a lot every day. Now, Dara still feels worthless and wants to sleep a lot. She continues to feel extremely sad and exhausted all day. (85 words)

Dara’s distressed behavioral vignette. For several months later, Dara sometimes became easily pessimistic when completing everyday tasks or interacting with the people she saw everyday. At such times she also felt a bit on edge, but that didn’t happen too often. Overall, she concentrated on the positive things in her life. A few times she overslept, but this didn’t affect her in any other way. Now, she generally feels satisfied overall and lives a productive life. She rarely oversleeps and she tries to be helpful towards friends and neighbors. (86 words)

Cathy’s “Recent Past” Vignettes.

Cathy’s traumatic event, intense reaction “recent past” vignette. Cathy is a physician who has been working in her town’s only hospital for 5 years. One morning, a chemical factory outside of town exploded. She heard it, felt the ground shake, and saw smoke in the direction of
the factory. She was at the hospital when it happened. This took her completely by surprise. She was shocked and felt her entire body freeze in horror. She couldn’t believe it. She had undergone disaster response training years ago, but had never imagined using those skills for this. She couldn’t believe this was happening to her right now. Within minutes the emergency room was full of people who had been at or near the explosion, and Cathy went to the emergency room to help. She couldn’t take up time and equipment helping one person who was very likely to die when she knew she could save another person if she acted quickly. This horrified Cathy; it completely went against everything she believed, and everything she knew. People’s families poured in, crying and screaming. She felt deeply rattled. How could anyone think in such a terrifying moment? She wasn’t sure if she could keep going. It was too horrific. She had to just let some people die. Cathy felt absolutely horror-stricken about this. She hated it, but knew deep down there wasn’t another option. She wanted to leave but she couldn’t; this was her obligation as a doctor. She noticed she was shaking. She tried to stop but couldn’t. She was terrified. She worked constantly for 24 hours before sleeping. The hospital was in chaos for days. (Event: 121 words, Reaction: 137 words)

Cathy’s traumatic event, mild reaction “recent past” vignette. Cathy is a physician who has been working in her town’s only hospital for 5 years. One morning, a chemical factory outside of town exploded. She heard it, felt the ground shake, and saw smoke in the direction of the factory. She was at the hospital when it happened. This was a bit jarring. Before this sudden event, it had simply seemed to be a morning like any other. She had undergone disaster response training years ago, but now she realized that she had never used that skill set. She frowned and wondered whether this situation could have been prevented. Within minutes the emergency room was full of people who had been at or near the explosion, and Cathy went to the emergency room
to help. She couldn’t take up time and equipment helping one person who was very likely to die when she knew she could save another person if she acted quickly. *This bothered Cathy; it went against her natural instincts. On the other hand, she understood its necessity.* People’s families poured in, crying and screaming. *She felt sympathetic; no one wanted to experience such a painful moment. She hoped every doctor would be able to keep going. It was fatiguing for her.*

She had to just let some people die. *This got to Cathy – she would prefer not to be in this situation. She kept thinking that no one can go back in time and change the bad things that happen, so good doctors are always needed.* She worked constantly for 24 hours before sleeping.

The hospital was in chaos for days. (Event: 121 words, Reaction: 137 words)

**Cathy’s everyday event, intense reaction “recent past” vignette.** Cathy is a physician who has been working in her town’s only hospital for 5 years. One morning, military officers came to the hospital to recruit physicians to become combat medics. They conveyed a medic’s life with photographs and stories. In one story, they described the emergency situation before and after a roadside bomb. *This took her completely by surprise. She was shocked and felt her entire body freeze in horror. She couldn’t believe it.* Cathy saw pictures of people who had been near the sudden explosion. *She had undergone disaster response training years ago, but had never imagined using those skills for this. She couldn’t believe this was happening to her right now.* The military officers described how medics cannot take up time and equipment helping one person who is likely to die when another person can be saved if they act quickly. *This horrified Cathy; it completely went against everything she believed, and everything she knew.* She saw photos of victims with severe injuries. *She felt deeply rattled. How could anyone think in such a terrifying moment? She wasn’t sure if she could keep going. It was too horrific.* The military officers explained that sometimes medics have to just let people die. *Cathy felt absolutely*
horror-stricken about this. She hated it, but knew deep down there wasn’t another option. She wanted to leave but she couldn’t; this was her obligation as a doctor. She noticed she was shaking. She tried to stop but couldn’t. She was terrified. Cathy has never been in the military, knows no one who has served. It is not the case that Cathy reacted this way because she was reminded of something in her past; she was truly this upset by the presentation. (Event: 121 words, Reaction: 164 words)

Cathy’s everyday event, mild reaction “recent past” vignette. Cathy is a physician who has been working in her town’s only hospital for 5 years. One morning, military officers came to the hospital to recruit physicians to become combat medics. They conveyed a medic’s life with photographs and stories. In one story, they described the emergency situation before and after a roadside bomb. This was a bit jarring. Before this sudden event, it had simply seemed to be a morning like any other. Cathy saw pictures of people who had been near the sudden explosion. She had undergone disaster response training years ago, but now she realized that she had never used that skill set. She frowned and wondered whether this situation could have been prevented. The military officers described how medics cannot take up time and equipment helping one person who is likely to die when another person can be saved if they act quickly. This bothered Cathy; it went against her natural instincts. On the other hand, she understood its necessity. She saw photos of victims with severe injuries. She felt sympathetic; no one wanted to experience such a painful moment. She hoped every doctor had been able to keep going. It would be fatiguing for her. The military officers explained that sometimes medics have to just let people die. This got to Cathy – she would prefer not to be in that situation. She kept thinking that no one can go back in time and change the bad things that happen, so good doctors are always needed.
Cathy has never been in the military and knows no one who has served. (Event: 121 words, Reaction: 137 words)

**Cathy’s Behavioral Vignettes: “Subsequent Behaviors”**

**Cathy’s PTSD behavioral vignette.** Over the following several months, Cathy felt constantly on alert for something bad to happen. She kept seeing images of that morning in her mind, and avoided other people who broached the subject because she feared reliving those memories. She lost interest in her hobbies and found it hard to concentrate. Her work suffered and she often left in the middle of the day. Now, months later, her mind is still plagued by images of that morning. (77 words)

**Cathy’s MDD behavioral vignette.** Over the following several months, Cathy felt increasingly empty inside. She cried a lot, and often woke up at odd hours, unable to get back to sleep. She found herself speaking slowly whenever she tried to describe her empty feeling and total lack of energy. Now, months later, she’s still struggling with a strong feeling of emptiness. She still cries a lot, and very often has trouble sleeping through the night. She feels lethargic and worthless. (76 words)

**Cathy’s distressed behavioral vignette.** Over the following several months, Cathy sometimes felt a bit on edge in the mornings during her first hour or so of work, but felt calm once she got into her usual routine. One evening, she became sad while thinking about cities abroad that had been recently attacked or bombed. On the whole, she continued to find her work enjoyable. Now, months later, she’s usually in good spirits. She feels generally content with her career and life. (77 words)

**Martin’s “Recent Past” Vignettes.**
Martin’s traumatic event, intense reaction “recent past” vignette. Martin was in a car accident about two years ago. He lives with his two kids. One day, he was driving everyone to the grocery store, just like he always did. As Martin was crossing an intersection (and he had the right of way), a tractor-trailer came out of nowhere and slammed into his car. As it happened, Martin was more afraid than he had ever felt before. He screamed and his heart rate skyrocketed. His car was crushed on the passenger side, where his daughter was in the backseat. The accident severely injured her. Martin was extremely afraid for his daughter’s life. He went into a state of panic. Inside, he was screaming in fear for her. She was rushed to the hospital and was declared to be in critical condition. Martin was too scared to even move. He felt dreadful and distressed. The accident was all he could think about. He felt like his entire world had screeched to a halt. His body was shaking and he broke out in sweat. Slowly, Martin’s daughter began to improve. He learned that she would make a full recovery after a lot of treatment. Martin and his son weren’t physically injured in the accident. That was absolutely the worst experience Martin believes he has ever had. He was completely overwhelmed by fear. He knew that it was an accident, but wished that it had never happened. (Event: 117 words, Reaction: 122 words)

Martin’s traumatic event, mild reaction “recent past” vignette. Martin was in a car accident about two years ago. He lives with his two kids. One day, he was driving everyone to the grocery store, just like he always did. As Martin was crossing an intersection (and he had the right of way), a tractor-trailer came out of nowhere and slammed into his car. His car was crushed on the passenger side, where his daughter was in the backseat. The accident severely injured her. As it happened, Martin was quite surprised and froze for a moment. He realized he was holding his breath, and exhaled. Martin was all right, but he was concerned about his daughter after the accident. He shook his head and wondered whether she would be okay – he
hoped so. She was rushed to the hospital and was declared to be in critical condition. Martin was taken off guard by the suddenness of it. He realized his heart was beating quickly. He was tense, but soon began to calm down. He thought about the driver who ran the stop sign; this accident could have been prevented. Slowly, Martin’s daughter began to improve. He learned that she would make a full recovery after a lot of treatment. Martin and his son weren’t physically injured in the accident. This was quite unexpected, and made Martin feel less complacent than usual. He decided to be more vigilant when driving on his regular routes. (Event: 117 words, Reaction: 122 words)

Martin’s everyday event, intense reaction “recent past” vignette. While he was driving in the city with his two kids, Martin witnessed a car accident at a 4-way stop. The car in front of him stopped at the stop sign, and Martin stopped just behind. As the front car pulled into the intersection, a driver travelling the opposite direction drove right past the stop sign and hit the car, right in front of Martin. As it happened, Martin was more afraid than he had ever felt before. He screamed and his heart rate skyrocketed. It was very loud and happened quickly. Martin was extremely afraid for those people’s lives. He went into a state of panic. Inside, he was screaming in fear for them. It turned out that everyone involved was unharmed, although upset, and both cars were damaged. Martin was too scared to even move. He felt dreadful and distressed. The accident was all he could think about. He felt like his entire world had screeched to a halt. His body was shaking and he broke out in sweat. Martin didn’t know either of the drivers, and didn’t interact with them. That was absolutely the worst experience Martin believes he has ever had. He was completely overwhelmed by fear. He knew that it was an accident, but wished that it had never happened. It is not the case that Martin reacted this way because he was reminded of
something in his past; he was truly this upset about the car accident alone. (Event: 111 words, Reaction: 151 words)

Martin’s everyday event, mild reaction “recent past” vignette. While he was driving in the city with his two kids, Martin witnessed a car accident at a 4-way stop. The car in front of him stopped at the stop sign, and Martin stopped just behind. As the front car pulled into the intersection, a driver travelling the opposite direction drove right past the stop sign and hit the car, right in front of Martin. It was very loud and happened quickly. As it happened, Martin was quite surprised and froze for a moment. He realized he was holding his breath, and exhaled. Martin was all right, but he was concerned about the people in the accident. He shook his head, and wondered whether they would be okay – he hoped so. It turned out that everyone involved was unharmed, although upset, and both cars were damaged. Martin was taken off guard by the suddenness of it. He realized that his heart was beating quickly. He was tense, but soon began to calm down. He thought about the driver who ran the stop sign; this accident could have been prevented. Martin didn’t know either of the drivers, and didn’t interact with them. He has never been in a car accident, and doesn’t know anyone who has been seriously injured in one. This was quite unexpected, and made Martin feel less complacent than usual. He decided to be more vigilant when driving on his regular routes. (Event: 111 words, Reaction: 122 words)

Martin’s Behavioral Vignettes: “Subsequent Behaviors”

Martin’s PTSD behavioral vignette. For a few months after, Martin refused to leave the house because of how dangerous he felt the world was after experiencing this firsthand. He had a hard time falling asleep, and suffered terrible nightmares about what happened. Martin also experienced extremely strong bursts of anger – too strong to contain. He knew that this strained his relationships, but couldn’t control himself. Now, he still won’t leave the house because of
how dangerous he believes the world is, and he has trouble sleeping. He still becomes angry often. (87 words)

**Martin’s MDD behavioral vignette.** For a few months after, Martin essentially stopped caring about his daily life, and found it impossible to concentrate on his work. His boss reprimanded him for missing project deadlines. Most of the time, he felt overwhelmingly guilty, and a few co-workers mentioned that he was losing a lot of weight. Now, Martin still feels uninterested in his daily life and continues to have difficulty focusing. He has noticed himself picking at his food rather than eating it. Martin still feels really guilty about what happened. (85 words)

**Martin’s distressed behavioral vignette.** For a few months after, Martin sometimes felt anxious and was very attentive driving on his regular routes, especially with his children in the car. He felt occasional twinges of worry, but they were minimal and dissipated quickly. Now, he rarely feels apprehensive or worried while driving, and feels calm in general. He still does feel occasional slight anxiety about his children’s safety, especially when driving with them around town or in the neighborhood, but not much. Martin enjoys his job and spending time with his children. (87 words)

**Justin’s “Recent Past” Vignettes.**

**Justin’s traumatic event, intense reaction “recent past” vignette.** Justin served in the military in Iraq. He loathed the violence that he witnessed there. *He felt a deep sense of dread about it.* On one occasion, when swept away by the group spirit, he killed a civilian when it was not necessary for his survival to do so. *When it happened, Justin felt an overwhelming surge of terror come over him. His head pounded and he felt like the world was rushing around him.* He was horrified. His actions seemed to be totally out of keeping with his character. *Justin felt his*
palms sweat. He thought that he had done a terrible thing, and realized that he could hurt someone at any moment. He was completely terrified by what he had done, and felt absolutely awful about that incident. Justin finished his tour of duty and was honorably discharged. After returning home, he was not able to find work. Instead, he lived solely on different forms of government assistance. (Event: 81 words, Reaction: 86 words)

Justin’s traumatic event, mild reaction “recent past” vignette. Justin served in the military in Iraq. He loathed the violence that he witnessed there. He preferred not to even think about it. On one occasion, when swept away by the group spirit, he killed a civilian when it was not necessary for his survival to do so. When it happened, Justin felt like this was an unfortunate part of war. He felt uncomfortable right then, as if he was too hot, and wished he hadn’t done it. His actions seemed to be totally out of keeping with his character. Justin thought that he was certainly not proud of this. He thought it was unlucky that he had been in that situation at all, knew he would probably never be in it again. He felt somewhat troubled about that incident. Justin finished his tour of duty and was honorably discharged. After returning home, he was not able to find work. Instead, he lived solely on different forms of government assistance. (Event: 81 words, Reaction: 86 words)

Justin’s everyday event, intense reaction “recent past” vignette. Justin recently bumped into his neighbor, Matt, at a neighborhood event. Matt had just returned from serving in the military in Iraq. Talking about this gave Justin a deep sense of dread. Matt said that on one occasion, when he was swept away by the group spirit, he killed a civilian when it was not necessary for his survival. When he heard this, Justin felt an overwhelming surge of terror come over him. His head pounded and he felt like the world was rushing around him. He was horrified. Matt said this was totally out of keeping with his character, and he became pretty upset
talking about it. Justin felt his palms sweat. He thought that Matt had done a terrible thing, and realized that he could hurt someone at any moment. He was completely terrified by what Matt had done, and felt absolutely awful about that story. Justin has never served in the military, and isn’t particularly close to Matt. It is not the case that Justin reacted this way because he was reminded of something in his past; he was truly this upset by his interaction with Matt. (Event: 74 words, Reaction: 115 words)

Justin’s everyday event, mild reaction “recent past” vignette. Justin recently bumped into his neighbor, Matt, at a neighborhood event. Matt had just returned from serving in the military in Iraq. Justin preferred not to even think about it. Matt said that on one occasion, when he was swept away by the group spirit, he killed a civilian when it was not necessary for his survival. When he heard this, Justin felt like this was an unfortunate part of war. He felt uncomfortable right then, as if he was too hot, and wished he hadn’t done it. Matt said this was totally out of keeping with his character, and he became pretty upset talking about it. Justin thought that Matt was certainly not proud of this. He thought it was unlucky that Matt had been in that situation at all, knew he would probably never be in it again. He felt somewhat troubled about that story. Justin has never served in the military, and isn’t particularly close to Matt. (Event: 74 words, Reaction: 86 words)

Justin’s Behavioral Vignettes: “Subsequent Behaviors”

Justin’s PTSD behavioral vignette. For some months afterward, Justin felt intense anxiety and his vision went hazy every time he saw military personnel. He couldn’t stop thinking about the terrors of war. He sped in his car everywhere he went, but he never felt like he was going fast enough. If anyone he knew talked about the military, he would snap at them and steer
clear of them for weeks. Now, about a full year later, he is constantly haunted by terrible thoughts of war. He always speeds even though he knows he really should not do so. (94 words)

**Justin’s MDD behavioral vignette.** For some months afterward, Justin kept experiencing a terribly sad mood. He felt like this almost constantly as he went through his day-to-day routine. He had a very hard time falling asleep. At times, Justin really couldn’t concentrate at all. He found himself eating a huge amount of food at every meal, but he never felt satisfied. Now, about a full year later, Justin is still almost always in a sad mood and has trouble concentrating. He craves large portions of food, but he never feels satisfied. Justin often cannot fall asleep. (92 words)

**Justin’s distressed behavioral vignette.** For some months afterward, Justin occasionally was antagonistic towards a neighbor or co-worker, but most of the time he was friendly. Occasionally everyday difficulties made him feel frustrated, but no more so than he supposed others felt with their own everyday difficulties. Justin almost always enjoyed spending time with his family and friends, talking and relaxing. Now, about a full year later, he rarely gets grouchy and is usually happy. He sometimes drinks with his friends, but it does not impact his work or relationships with friends or family. Justin generally feels satisfied with his life. (96 words)

**Lucy’s “Recent Past” Vignettes.**

**Lucy’s traumatic event, intense reaction “recent past” vignette.** Lucy teaches at an elementary school where there was a shooting several months ago. She was with her class of children in the playground when they were pinned down by sniper fire. *Immobilized and vulnerable, she became completely paralyzed with fear. She felt like they were all going to die. Lucy felt complete and total terror and could barely breathe. Her thoughts raced and she prayed for it to stop.* The shooting lasted for 15 minutes. The sniper killed one student and injured many
more. Lucy tried to hide behind a trash can; from there she saw a child get shot and fall to the ground. This absolutely crushed her. She felt shocked, helpless, and ineffective. After the gunfire ceased, no one moved until the police stormed the sniper’s apartment, only to find that he had killed himself. She was not relieved. She realized that even the seemingly safest places, like an elementary school, could easily be attacked at any time. She was not able to help anyone because she could have been shot. As the event unfolded, she experienced an extremely heightened sense of fear, and felt like she had failed her students. She couldn’t believe that something so terrible had happened to her at work. (Event: 102 words, Reaction: 108 words)

Lucy’s traumatic event, mild reaction “recent past” vignette. Lucy teaches at an elementary school where there was a shooting several months ago. She was with her class of children in the playground when they were pinned down by sniper fire. Lucy was caught off guard and confused about what was going to happen. She tried to be confident for the children. She couldn’t think clearly, because it was too loud. Lucy’s head hurt. She hoped it would end soon. The shooting lasted for 15 minutes. The sniper killed one student and injured many more. Lucy tried to hide behind a trash can; from there she saw a child get shot and fall to the ground. She felt that this couldn’t be helped, and just tried to be calm and patient. After the gunfire ceased, no one moved until the police stormed the sniper’s apartment, only to find that he had killed himself. She was relieved. She realized that even the seemingly safest places, like a school, could be vulnerable if targeted. She was not able to help anyone because she could have been shot. She came away from this feeling quite ruffled, but certain that she had done the best she could at the time. She was very surprised to have had this experience. (Event: 102 words, Reaction: 107 words)
Lucy’s everyday event, intense reaction “recent past” vignette. Lucy teaches at an elementary school where there was a lockdown drill several months ago. She hadn’t known about the drill in advance because she had been out sick and missed the memo. She was with her class of children when the lockdown siren began. Lucy and her students all crouched under their desks. Immobilized and vulnerable, she became completely paralyzed with fear. She felt like they were all going to die. Lucy felt complete and total terror and could barely breathe. Her thoughts raced and she prayed for it to stop. She tried to calm her students down, but some of the children seemed pretty scared. This absolutely crushed her. She felt shocked, helpless, and ineffective. The drill lasted for 15 minutes. Three emergency service personnel outfitted with emergency equipment visited the classroom. After the siren ceased, no one moved until a police officer returned to say that it was over. She was not relieved. She realized that even the seemingly safest places, like an elementary school, could easily be attacked at any time. As the event unfolded, she experienced an extremely heightened sense of fear, and felt like she had failed her students. She couldn’t believe that something so terrible had happened to her at work. It is not the case that Lucy reacted this way because she was reminded of something in her past; she was truly this upset by the drill alone. (Event: 102 words, Reaction: 136 words)

Lucy’s everyday event, mild reaction “recent past” vignette. Lucy teaches at an elementary school where there was a lockdown drill several months ago. She hadn’t known about the drill in advance because she had been out sick and missed the memo. She was with her class of children when the lockdown siren began. Lucy and her students all crouched under their desks. Lucy was caught off guard and confused about what was going to happen. She tried to be confident for the children. She couldn’t think clearly, because it was too loud. Lucy’s head hurt. She hoped it would end soon. She tried to calm her students down, but some of the children
seemed pretty scared. *She felt that this couldn't be helped, and just tried to be calm and patient.*

The drill lasted for 15 minutes. Three emergency service personnel outfitted with emergency equipment visited the classroom. After the siren ceased, no one moved until a police officer returned to say that it was over. *She was relieved. She realized that even the seemingly safest places, like a school, could be vulnerable if targeted. She came away from this feeling quite ruffled, but certain that she had done the best she could at the time. She was very surprised to have had this experience.* (Event: 102 words, Reaction: 107 words)

**Lucy’s Behavioral Vignettes: “Subsequent Behaviors”**

**Lucy’s PTSD behavioral vignette.** Soon afterward, Lucy felt plagued by this event. She thought about it constantly – in great detail as if it were reoccurring. She was never in a good mood, and had a short temper with friends and co-workers. She also frequently woke up in the middle of the night, wracked with shame. She did everything she could to avoid talking about it. Now, it has been a few months, and she still feels ashamed and plagued by thoughts of the event, even when she tries not to think about it. She tries not to talk about it. (96 words)

**Lucy’s MDD behavioral vignette.** Soon afterward, Lucy started to experience an overwhelming feeling of hopelessness that persisted all day and night. She felt extremely lethargic, as if she needed 10 times the usual amount of energy to complete everyday tasks. She also found it very difficult to concentrate, which negatively impacted her work. She lost pleasure in things she once enjoyed. Now, it has been a few months, and she still feels overwhelmingly hopeless, she has very little energy, and no ability to concentrate. Her effectiveness at work has suffered a lot, and sometimes her friends comment that she never looks happy. (98 words)

**Lucy’s distressed behavioral vignette.** Soon afterward, Lucy felt uneasy but went on with her everyday life. Once, she woke up in the middle of the night feeling wide-awake and a
little bit tense, but was able to fall back to sleep. On days when she felt irritable, she talked with friends and colleagues about what was bothering her. This usually helped her feel better. Now, it has been a few months, and she occasionally feels a little uneasy; however, this doesn’t significantly interfere with her life. Sometimes she feels a bit irritable, but generally she’s calm and believes she is okay. (97 words)

Diana’s “Recent Past” Vignettes.

Diana’s traumatic event, intense reaction “recent past” vignette. One day, months ago, Diana was in the midst of coordinating the world premiere of a new ballet. Construction problems had made it uncertain whether the opening deadline would even be met. The set designer was volatile, and threatened to walk out on the project altogether. This made Diana feel terribly powerless. She worried about what might go wrong next. She noticed herself taking many deep breaths. While Diana was dealing with all of this, she also had heavier responsibilities at home because both of her children came down with the flu. This made Diana feel like she could do little without help. It felt like everything she tried to do was inevitably difficult. Then, she received word that her best friend was badly injured in a car crash. It left her in a coma. Diana couldn’t understand this, and felt frozen in a state of utter helplessness. Then she broke down crying and sobbing. This was too much for her to bear because it was her best friend. This was also too much for her on top of everything else. Diana felt like she had lost all of her ability to do anything right in her life. She felt desperate and powerless. Diana is an only child, and has been very close to her best friend ever since grade school. People would often comment that they were like sisters. After all this, Diana felt incredibly weak. (Event: 105 words, Reaction: 122 words)
Diana’s traumatic event, mild reaction “recent past” vignette. One day, months ago, Diana was in the midst of coordinating the world premiere of a new ballet. Construction problems had made it uncertain whether the opening deadline would even be met. The set designer was volatile, and threatened to walk out on the project altogether. Diana remembered that sometimes people are not particularly considerate or helpful. She tried to shrug it off, but was annoyed. While Diana was dealing with all of this, she also had heavier responsibilities at home because both of her children came down with the flu. This made Diana realize how much we rely on others. If she took other people for granted, did they do the same to her? Then, she received word that her best friend was badly injured in a car crash. It left her in a coma. This was unfortunate, and Diana wished she could do something. She was bothered by this turn of events, but was so busy that she hardly had time to think about it. She talked with others about what had happened, which reassured her and helped put her mind at ease. She tried not to let the accident distract her from the premiere, but had trouble enjoying herself. Diana is an only child, and has been very close to her best friend ever since grade school. People would often comment that they were like sisters. After all this, Diana felt pretty tired. (Event: 105 words, Reaction: 122 words)

Diana’s everyday event, intense reaction “recent past” vignette. One day, months ago, Diana was in the midst of organizing a dinner party. One friend called and wanted her to change the date last-minute, but she couldn’t because the scheduled night worked best for everyone else. This made Diana feel terribly powerless. She worried about what might go wrong next. She noticed herself taking many deep breaths. On that day, she also had to drive her children to their dance lessons, because the family friend who usually drives them was away that week. This made Diana feel like she could do little without help. It felt like everything she tried to do was inevitably difficult. Then, she received word that her best friend was in a fender bender just
before the party. She didn’t need to go to the hospital, but didn’t feel like coming to the party anymore. Diana couldn’t understand this, and felt frozen in a state of utter helplessness. Then she broke down crying and sobbing. This was too much for her to bear because it was her best friend. This was also too much for her on top of everything else. Diana felt like she had lost all of her ability to do anything right in her life. She felt desperate and powerless. Her friend wanted them to go ahead, so Diana did. The party went smoothly even though a few people missed it. After all this, Diana felt incredibly weak. It is not the case that Lucy reacted this way because she was reminded of something in her past; she was truly this upset about her friend alone.

(Event: 108 words, Reaction: 150 words)

Diana’s everyday event, mild reaction “recent past” vignette. One day, months ago, Diana was in the midst of organizing a dinner party. One friend called and wanted her to change the date last-minute, but she couldn’t because the scheduled night worked best for everyone else. Diana remembered that sometimes people are not particularly considerate or helpful. She tried to shrug it off, but was annoyed. On that day, she also had to drive her children to their dance lessons, because the family friend who usually drives them was away that week. This made Diana realize how much we rely on others. If she took other people for granted, did they do the same to her? Then, she received word that her best friend was in a fender bender just before the party. She didn’t need to go to the hospital, but didn’t feel like coming to the party anymore. This was unfortunate, and Diana wished she could do something. Her friend wanted them to go ahead, so Diana did. She was bothered by this turn of events, but was so busy that she hardly had time to think about it. She talked with others about what had happened, which reassured her and helped put her mind at ease. She tried not to let the accident distract her from the party, but
had trouble enjoying herself. The party went smoothly even though a few people missed it. After all this, Diana felt pretty tired. (Event: 108 words, Reaction: 122 words)

Diana’s Behavioral Vignettes: “Subsequent Behaviors”

Diana’s PTSD behavioral vignette. Over the next couple of months, Diana felt as if she had completely lost her ability to feel happiness. She couldn’t stop thinking back to that day in her mind – she felt like it was happening over and over again. She began spending money on a lot of things she didn’t need, always putting it on her credit card. She still hasn’t spoken to anyone about that day and doesn’t intend to do so. Now, she still thinks about it constantly. She has maxed out her credit cards with impulse buys. She never feels happy. (95 words)

Diana’s MDD behavioral vignette. Over the next couple of months, Diana slowly gave up all of her responsibilities until she had very few left. She slept for 12 hours each night, sometimes even more. Many days, she did not even get out of bed during the day; she just wanted to stay there. She felt constantly very sad inside. Now, she still has not taken on any new projects. She spends many days in bed, sleeping on and off, and feeling very down. Generally, she thinks that friendship efforts are futile. She spends most of her time alone. (94 words)

Diana’s distressed behavioral vignette. Over the next couple of months, Diana tried not to take on any new responsibilities and offered them to a colleague instead. Some days, she stayed at home and took care of her responsibilities from there. If she felt down, she spent time playing with her children, which almost always improved her mood. Whenever they had time, Diana and her husband would spend time together discussing silly things instead of talking about everyday stressors. Now, Diana is involved in several community activities. She continues to spend time with close friends. She feels satisfied with her life. (96 words)
Appendix B

College Student Open-Ended Disorder Identification Responses

In this section, we describe more fully all of the open-ended disorder identification responses made by college students. As a reminder, each participant read two vignettes describing PTSD, two MDD, and two distressed behaviors. For each vignette, they were asked to type in up to three disorder identifications. They were explicitly instructed that one disorder identification was required for each vignette. They were also asked to rate the likelihood of each disorder identification on a scale of 1-9, where 1=very unlikely and 9=very likely (please see “College Students’ Identification of Mental Disorders in Context,” Table A2 for exact text). If they felt that no disorder identification was needed, they were instructed to type “none,” and to rate the likelihood of this answer as well.

Eighty people saw six vignettes each for a total of 480 vignettes in this study. One hundred and sixty vignettes described PTSD behaviors, 160 described MDD behaviors, and 160 described distressed behaviors. For all 480 vignettes, people offered a total of 644 disorder identifications, 231 for PTSD behaviors, 224 for MDD behaviors, and 189 for distressed behaviors. On average, people made 1.34 disorder identification per vignette. Thus, people were willing to make more than one disorder identification for a single vignette.

Below, we describe and discuss the DSM disorder identifications college students made, the occasions on which they indicated no disorder identification was necessary, and the non-DSM disorder identification responses they made. Next, we describe and discuss their indications of multiple disorder identifications for a single vignette and the influence of context.

DSM Disorder Identifications
College students demonstrated knowledge of a wide variety of actual disorders. They identified a *DSM-IV-TR* (APA, 2000) or *DSM-5* (APA, 2013) disorder 74% of the time (476 out of the 644 total responses). The disorder identifications included PTSD (142), Depression (172), MDD (2), Anxiety Disorder (82), Paranoia (Paranoid Personality Disorder, 12), Generalized Anxiety Disorder (7), Phobia (Specific Phobia, 7), Obsessive Compulsive Disorder (7), Mood Disorder (6), Panic Disorder (5), Bipolar Disorder (4), Eating Disorder (4), Attention-Deficit/Hyperactivity Disorder (3), Insomnia Disorder (3), Acute Stress Disorder (3), Social Anxiety Disorder (3), Agoraphobia (3), Schizophrenia (3), Alcoholism / Alcohol Abuse / Alcohol Addiction (Alcohol Use Disorder, 3), Binge Eating Disorder (2), Other Personality Disorder (2), Dissociative Disorder (1), and Borderline Personality Disorder (1). Please see “College Students’ Identification of Mental Disorders in Context,” Table A3 for details by behavioral vignette.

Overall, college students identified 22 DSM disorders. The large number of DSM disorder identifications reveals that people have at least heard of, and can attempt to apply, a range of mental disorder diagnoses used in professional clinical practice. This may reflect a more general phenomenon that college students map their mental disorder concepts to those used by professionals (Shaw, 2002).

College students made the “appropriate” disorder identifications for 62% of vignettes. That is, they indicated the very disorder that we described (as carefully as possible) in the vignette. Again, each vignette fulfilled all of the diagnostic criteria for either a PTSD diagnosis, a MDD diagnosis, or no disorder at all. Overall, college students indicated PTSD for 48% of PTSD behavioral vignettes, depression or MDD for 73% of MDD behavioral vignettes, and indicated “none” for 64% of distressed behavioral vignettes. Thus, college students were able to recognize the disordered behavioral symptoms in the vignettes as a particular clinical disorder,
demonstrating their knowledge of PTSD and MDD. Again, this may reflect a phenomenon wherein college students accurately incorporate professional clinical concepts of mental disorders into their own understanding and discourse of abnormal behavioral symptoms in themselves and others, perhaps in an attempt to facilitate communication with professionals if and when they have contact (Shaw, 2002).

**Multiple DSM disorder identifications for single vignettes.** College students indicated two or three DSM disorder identifications for 25% of the total vignettes. Fifty-one of the PTSD behavioral vignettes (32%), 47 of the MDD behavioral vignettes (29%), and 23 of the distressed behavioral vignettes (14%) received two or three DSM disorder identifications. It is possible that at least some of these college students thought that the hypothetical people described in the vignettes were comorbid. In reality, PTSD is highly comorbid with other DSM disorders. People with DSM-5 (APA, 2013) PTSD symptoms are 80% more likely than those without PTSD symptoms to qualify for another DSM-5 diagnosis (APA, 2013). Disorders comorbid with PTSD include depressive disorders, anxiety disorders, and substance abuse disorders. MDD also has a high comorbidity rate (estimated as high as 40%; Hasin, Goodwin, Stinson, & Grant, 2005) with substance abuse disorders, anxiety disorders, borderline personality disorder, and PTSD (to name a few). It may be the case that the student participants were aware of the high comorbidity of PTSD and MDD with other disorders, and this drove them to indicate multiple possible disorders for these vignettes. On the other hand, we gave people space to make up to three disorder identifications; they may have been influenced by demand characteristics to search for and indicate more than one disorder. Yet, multiple disorder identifications accounted for only one quarter of the sample, so it is unlikely that all participants felt compelled to make multiple disorder identifications for each vignette. In addition, college students could have interpreted the
task to mean that any second or third disorder identifications they offered for a single vignette were alternative possibilities. Finally, college students may hold a general notion that disorders don’t always occur in isolation, which may help explain why some distressed behaviors also received multiple DSM disorder identifications. College students identified 38% of people in vignettes with merely distressed behaviors as meriting various DSM disorder identifications, suggesting a possible tendency to generally over diagnose.

“No Diagnosis” Responses

College students demonstrated a general willingness to indicate that no disorder was present (“none”); 120 responses consisted of a “none” response (19% of total responses), seven for PTSD behaviors (3% of responses for PTSD behaviors), 10 for MDD behaviors (4%), and 103 for distressed behaviors (55%). The vast majority of “none” responses were made “appropriately,” that is, when the vignette described non-disordered behaviors. Thus, college students were not simply providing disorder identifications for every single vignette. Rather, they were also assessing behavioral vignettes for whether a disorder was present.

Non-DSM Disorder Identifications

The remaining 7% of disorder identification responses were not DSM disorders. People made a total of 47 such responses; 23 for PTSD behaviors (10% of responses for PTSD behaviors), 11 for MDD behaviors (5%), and 13 for distressed behaviors (7%). Interestingly, there were 13 trauma-related responses: survivor’s guilt (4), trauma (2), very mild PTSD (2), “some form of PTSD?” (1), vicarious PTSD (1), PTSD – shared (1), World-Terror Disorder (1), and extreme paranoia of the military (1). The non-disordered diagnosis responses further included 8 that resembled a DSM disorder, proposed DSM disorder, or symptom from a DSM disorder other than a trauma-related disorder: compulsive buying disorder (2), emotional
detachment (2), anger management disorder (1), “something with paranoia” (1), control freak (1), delusions (1). Nine responses did not approximate a known diagnosis: confident disorder (1), false shame (1), fear of responsibility (1), hyper empathy (1), repression (1), sequela (1), “inability to cope with perceived failure” (1), “overeating to compensate for his depression” (1), and multiple non-DSM disorder responses in one response box (1). People also made 14 other responses that indicated a state: stress (3), nervousness (3), helplessness (2), scared (1), shock (1), happiness (1), mild depression (1), slight anxiety (1), and agonal state (1). Finally, three responses admitted a lack of knowledge: “I don’t know” (2), and “Something. Not sure.” (1).

It is possible that college students believed the non-DSM responses were true disorders. Indeed, as noted above, quite a few of these responses approximated a DSM disorder or symptom (e.g., emotional detachment), but not strongly enough to be counted as a DSM disorder for the purpose of our analyses. College students may have been trying to describe a DSM disorder with which they were familiar, but were unable to retrieve the name at the time of the study. A second possibility is that college students lacked knowledge of DSM disorders, but were still attempting to complete the task by providing “diagnoses” as a clinician would. Thus, the demands of the task may have led people to elicit these types of DSM-like responses, even though they had poor knowledge of DSM disorders. Third, college students might not have been trying to approximate DSM disorders in their responses at all. Rather, people may simply have done their best to indicate a condition that they felt the person in the vignette possessed, using a label they found descriptive. Any of these three explanations might apply to a subset of the responses we obtained. These possibilities do not exhaust the possible reasons why college students made non-DSM disorder identifications in this study.

The Influence of Context on Disorder Identifications: Secondary Analyses
We ran chi square tests to examine the effect of event context (traumatic event, everyday event) and reaction context (intense reaction, mild reaction) for all three behavior types (PTSD, MDD, distressed) on the number of the following types of disorder identifications: anxiety disorder and all DSM disorder identifications combined (these were the only responses given by at least ten college students per behavior type). The statistical tests of primary theoretical interest (i.e., on PTSD identifications for PTSD behaviors, depression or MDD identifications for MDD behaviors, and “none” responses for distressed behaviors) are reported in full in the main study. Here, we report the secondary analyses in case they are of interest.

**Anxiety disorder identifications in context.** People indicated “anxiety disorder” for PTSD behaviors (N=34), MDD behaviors (N=21) and distressed behaviors (N=27). When people identified PTSD or distressed behaviors as cases of anxiety disorder, they did not do so differentially by event or reaction context (all ps > .271). Also, the test of event context was not significant for MDD behaviors (p = .579). However, the test of reaction type on anxiety disorder identifications was significant for MDD behaviors ($X^2[2, N = 80] = 5.17, p = .076$), such that people were more likely to identify an anxiety disorder for MDD behaviors after a severe reaction ($N = 12$) than a mild reaction ($N = 9$). Unfortunately, the difference in number of responses between conditions is so small that a clear interpretation of this finding is difficult. However, as slightly more anxiety disorder identifications occurred for MDD behaviors in severe than in mild reaction context, this might suggest that people’s notions of anxiety disorders include internal reactions. Indeed, a fear response is a common symptom of anxiety disorders (DSM-5, APA, 2013).

**All DSM disorder identifications in context.** We ran analyses to understand whether people made any type of DSM disorder identification more often in event or reaction context.
People indicated a *DSM* disorder, including the diagnoses of primary interest (e.g., PTSD diagnoses for PTSD behaviors), a total of 476 times combined, 200 for PTSD behaviors, 203 for MDD behaviors, and 73 for distressed behaviors. When PTSD or distressed behaviors were identified as a case of any *DSM* disorder, they were not identified differently in either event or reaction context (all $ps \geq .168$). However, the test of event type for any *DSM* disorder identification was significant for MDD behaviors, such that college students were marginally more likely to identify MDD behaviors as any *DSM* disorder in traumatic event context ($N=108$) compared to everyday event context ($N=75$; $X^2[5, N = 80] = 10.33, p = .066$). The test of reaction type for MDD behaviors was not significant ($p = .536$). Previous tests revealed that MDD behaviors were identified more often as PTSD in traumatic event context, but were not identified as cases of depression more or less often across contexts. Thus, we were interested in whether the above effect was driven by PTSD identifications for MDD behaviors. We ran two chi square tests of event and reaction context on all *DSM* identifications except for PTSD identifications, for MDD behaviors specifically. These tests revealed no significant effects (all $ps \geq .676$). Thus, PTSD diagnoses for MDD behaviors did drive this effect. People made all disorder identification decisions for MDD behaviors regardless of context, except for PTSD identifications for MDD behaviors, which they made more often in traumatic event context.

In conclusion, college students made many disorder identification responses, the majority were *DSM-IV-TR* (APA, 2000) and *DSM-5* (APA, 2013) disorders, and PTSD was consistently identified in traumatic event context.
Appendix C

Pilot Study: College Students’ Identification of Mental Disorders Using a Closed-Ended Measure of Diagnosis Judgments

In an initial pilot study, we also examined whether and how college students can identify PTSD and MDD in context. This was a test of the same hypotheses as in the main study, with a more constrained disorder identification measure for better control over the variability of disorder identification responses.

Methods

The Northeastern University Internal Review Board approved all experiment protocols.

Participants

We recruited 30 undergraduate students (27 female) at Northeastern University with no clinical training or experience, who participated for course credit. In our sample, participants reported a mean age of 18.5 years (range: 18-21), and self-identified as White (70%), Asian (23%), and Black (7%); none self-identified as Hispanic.

Materials and Procedure

We employed the same materials and design as in the main study with college students, with one exception. In place of the open-ended disorder identification measure, we asked college students to rate the likelihood that the person in the vignette has PTSD, and separately, the likelihood that the person in the vignette has MDD, using two closed-ended identification questions. The PTSD closed-ended identification question asked, “Given [X]’s subsequent behaviors, is [X] likely to have Post-Traumatic Stress Disorder?” The MDD closed-ended identification question asked, “Given [X]’s subsequent behaviors, is [X] likely to have Major Depressive Disorder?” Both questions were followed by the same instructions: “Using the scale
below, please rate the likelihood of this diagnosis on a scale of 1-9, where 1=very unlikely and 9=very likely.” For each vignette, college students also answered the psychological abnormality question and the difficulty understanding questions from the main study.

College students answered the same dependent variable for all six vignettes before moving on to answer another dependent variable for all six vignettes, until all four dependent variables had been presented. The order of dependent variables was randomized, and the order of vignettes presented for each dependent variable was also randomized.

**Results and Discussion**

Just as in the main study, we had two primary questions. First, we asked, are college students influenced by context in identifying mental disorders? Second, we asked, is college students’ ability to identify mental disorders systematically influenced only by traumatic events, only intense reactions, only the presence of both traumatic events and intense reactions, or none of these contexts? To do so, we examined whether event and reaction context affected the following disorder identifications: (1) people’s judgments of PTSD likelihood for PTSD behaviors, (2) people’s judgments of MDD likelihood for MDD behaviors, (3) people’s judgments of PTSD likelihood for distressed behaviors, and (4) people’s judgments of MDD likelihood for distressed behaviors. In each case, the critical comparisons were between the different event and reaction context conditions for the same behaviors. We simultaneously explored whether, and how, context affects judgments of psychological abnormality and difficulty understanding for PTSD behaviors, MDD behaviors, and distressed behaviors. For each dependent variable, we ran a 2 (Event Type) x 2 (Reaction Type) ANOVA. In all analyses, ratings for individual vignettes of the same behavior type were collapsed. All analyses were conducted at the $\alpha = .05$ level, and paired comparisons were Bonferroni-corrected.
Primary Analyses: Closed-Ended Disorder Identifications in Context

For the primary judgments of disorder identification likelihood, we expected to replicate the findings of the main study. That is, we expected college students’ judgments of the likelihood of PTSD for PTSD behaviors to be higher in traumatic event context than in everyday event context, and to be unaffected by reaction context. We also expected college students’ judgments of the likelihood of MDD for MDD behaviors to be unaffected by either event or reaction context. As described in Appendix B, college students identified distressed behaviors as having PTSD more often in traumatic event context compared to everyday event context (although marginally), and we expected to replicate this finding in likelihood judgments. Finally, we expected that MDD identifications for distressed behaviors would not be influenced by context, as in the main study.

PTSD disorder identification for PTSD behaviors. Just as in the main study, college students judged PTSD to be more likely for PTSD behaviors in the context of a traumatic event ($M = 7.71, SD = .96$), than in the context of an everyday event ($M = 5.80, SD = 2.37; F[1,26] = 7.70, p = .010, \eta^2_p = .23$). Further, there was no main effect of reaction ($p = .847, \eta^2_p < .01$); nor was there an interaction ($p = .969, \eta^2_p < .01$). Thus, college students’ judgments of PTSD likelihood in this pilot study aligned with what we found in the main study; that is, they cohered with the DSM-5’s requirement that PTSD be diagnosed only after an objectively traumatic event.

MDD disorder identification for MDD behaviors. Just as in the main study, college students did not judge MDD to be any more or less likely for MDD behaviors in context. That is, there were no main effects of event or reaction type and no interaction (all $ps \geq .464, \eta^2_p \leq .02$). Thus, although the context of a traumatic event led to increased PTSD identifications for PTSD, traumatic event context did not affect MDD identifications for MDD behaviors. Perhaps the
MDD behaviors in the vignette were sufficiently recognizable to college students as indicative of MDD regardless of context type. Interestingly, college students’ judgments adhered to DSM-5 (APA, 2013) criteria for MDD diagnosis, as we found in the main study.

**PTSD disorder identification for distressed behaviors.** For judgments of PTSD likelihood for distressed behaviors, we replicated the findings of our main study. There was a main effect of event, \( F[1,26] = 6.03, p = .021, \eta_p^2 = .19 \), such that college students judged PTSD to be more likely for *distressed, non-disordered* behaviors in the context of a traumatic event \( M = 3.70, SD = 1.84 \), than in the context of an everyday event \( M = 2.27, SD = 1.27 \). There was no main effect of reaction \( (p = .241, \eta_p^2 = .05) \) and no interaction \( (p = .994, \eta_p^2 < .01) \). Thus, lay people considered traumatic event context sufficient to indicate PTSD for behaviors that were merely distressed.

**MDD disorder identification for distressed behaviors.** Unlike in the main study, MDD identifications for merely *distressed* behaviors were influenced by context. There was a main effect of event, \( F[1,26] = 4.51, p = .043, \eta_p^2 = .15 \), such that college students judged MDD to be more likely for distressed behaviors in the context of a traumatic event \( M = 3.30, SD = 1.41 \), than in an everyday event \( M = 2.23, SD = 1.43 \). There was also a main effect of reaction, \( F[1,26] = 4.67, p = .040, \eta_p^2 = .15 \) such that college students judged MDD to be more likely for distressed behaviors in the context of an intense reaction \( M = 3.32, SD = 1.53 \), than a mild reaction \( M = 2.28, SD = 1.33 \). There was no interaction \( (p = .179, \eta_p^2 = .07) \). Thus, both traumatic event and intense reaction contexts independently increased MDD identifications for distressed behaviors relative to their milder counterparts. We might not have replicated the main study findings because college students might have perceived the distressed behaviors themselves to be *somewhat* indicative of MDD, and made their likelihood judgments
accordingly. Such perceptions then might have been further increased by context, because
college students may feel that depression typically occurs as a response to some type of stressful
event or intense reaction, leading them to judge MDD to be more likely for non-disordered
behaviors in such contexts.

Secondary Analyses: Judgments of Psychological Abnormality and Difficulty
Understanding

For the secondary judgments of psychological abnormality and difficulty understanding, we expected to fully replicate the findings in the main study, as these were the same vignettes employed in the main study, and the same dependent variables. That is, we expected college students to find PTSD and MDD behaviors to be less psychologically abnormal and less difficult to understand in traumatic event context, compared to everyday event context, and regardless of reaction context. Further, we expected that college students would not be influenced by context when making these judgments for distressed behaviors, as in the main study.

Psychological abnormality judgments for PTSD behaviors. In concert with our findings in the main study, college students found PTSD behaviors to be marginally less psychologically abnormal in the presence of a traumatic event ($M = 5.07, SD = 2.25$) than an everyday event ($M = 6.53, SD = 2.03$; $F[1,26] = 3.22, p = .085, \eta_p^2 = .11$). There was no main effect of reaction type ($p = .719, \eta_p^2 < .01$) or interaction ($p = .836, \eta_p^2 < .01$). Thus, the presence of a traumatic event tended to normalize students’ perceptions of PTSD behaviors.

Difficulty understanding judgments for PTSD behaviors. For difficulty understanding judgments for PTSD behaviors, there was a main effect of event type as in the main study ($F[1,26] = 6.95, p = .014, \eta_p^2 = .21$), such that people judged PTSD behaviors to be less difficult to understand in the presence of a traumatic event ($M = 3.27, SD = 1.67$) than an everyday event
There was also a marginally significant main effect of reaction type \( (F[1,26] = 3.38, p = .077, \eta_p^2 = .12) \), such that people judged PTSD behaviors to be less difficult to understand in the presence of an intense reaction \( (M = 3.46, SD = 1.63) \) than a mild reaction \( (M = 4.66, SD = 2.17) \). There was no interaction \( (p = .327, \eta_p^2 = .04) \). Thus, the presence of a traumatic event made PTSD behaviors seem less difficult to understand than did an everyday event, as in the main study. However, the context of an intense reaction also made PTSD behaviors seem less difficult to understand than did a mild reaction. Of note, both contexts influenced perceptions of difficulty understanding in the same direction. That is, college students seemed to be influenced by both types of context in making sense of PTSD behaviors, but there was no additive effect of the two types of context (i.e., there was no interaction).

**Psychological abnormality judgments for MDD behaviors.** Unlike in the main study, where traumatic event context made MDD behaviors seem less psychologically abnormal, there were no main effects, nor was there an interaction (all \( ps \geq .560; \text{all } \eta_p^2 \leq .01 \)). That is, context did not influence how psychologically abnormal MDD behaviors seemed in this study. This finding aligns somewhat with MDD disorder identifications obtained in this study and the main study, as these judgments were not influenced by context.

**Difficulty understanding judgments for MDD behaviors.** Just as we found in the main study, college students found MDD behaviors to be less difficult to understand in traumatic event context \( (M = 2.64, SD = 1.10) \) than in everyday event context \( (M = 4.67, SD = 1.96; \text{ F}[1,26] = 11.79, p = .002, \eta_p^2 = .31) \). There was no main effect of reaction type \( (p = .527, \eta_p^2 = .02) \), nor an interaction \( (p = .435, \eta_p^2 = .02) \). Thus, lay people’s perceptions of the difficulty understanding MDD behaviors were indeed affected by event context, even if perceptions of psychological abnormality were not.
Psychological abnormality judgments for distressed behaviors. Unlike in the main study, for psychological abnormality judgments for distressed behaviors, there was a main effect of event type. However, it was qualified by an interaction of event and reaction type (F[1,26] = 5.12, p = .032, \( \eta^2_p = .17 \)). There was no main effect of reaction type. Bonferroni-corrected independent samples t-tests were conducted to examine the interaction. In traumatic event context, the two reaction contexts did not affect judgments differently (p = .554). However, in everyday event context, distressed behaviors seemed less psychologically abnormal after a mild reaction (\( M = 1.63, SD = .58 \)) than after a severe reaction (\( M = 3.36, SD = 1.31; t[13] = 3.38, p = .005 \)). Furthermore, in severe reaction context, the two event types did not affect judgments differently (p = .920). In mild reaction context, however, distressed behaviors seemed less psychologically abnormal after an everyday event (\( M = 1.63, SD = .58 \)) than after a traumatic event (\( M = 3.88, SD = 1.89; t[14] = 3.23, p = .022 \)). Generally, people rated distressed behaviors, which are mildly negative, to be less psychologically abnormal in the context of a proportionately mildly negative combinations of an everyday event paired with a mild reaction, compared to all three other event and reaction combinations (e.g., a traumatic event paired with a severe reaction).

Difficulty understanding judgments for distressed behaviors. For difficulty understanding judgments for distressed behaviors, there were no main effects or interactions (all \( p \text{'s} \geq .267; \text{all } \eta^2_p \leq .05 \)), just as in the main study.

Summary of Pilot Study Findings

In this pilot, we replicated the key findings from the main study, including the main disorder identification responses of interest. First, we found that college students rated PTSD to be more likely for PTSD behaviors in traumatic event context than in everyday event context.
using the closed-ended measure, just as we found with PTSD identifications in the main study with the open-ended measure. Second, we found that college students were not influenced by context when rating the likelihood of MDD for MDD behaviors, just as we found with MDD identifications in the main study with the open-ended measure. Third, we found that college students rated PTSD to be more likely for distressed behaviors in traumatic event context than in everyday event context using the closed-ended measure, just as we found with PTSD identifications in the main study with the open-ended measure. For the secondary analyses, we replicated main study findings of psychological abnormality judgments for PTSD behaviors and difficulty understanding judgments for MDD and distressed behaviors.

**Implications and Limitations**

The results of this study are informative because they demonstrate that the closed-ended disorder identification measure is comparable to the open-ended disorder identification measure, when assessing the influence of context on PTSD and MDD identifications. Overall, the use of the closed-ended measure in this study was an important first step, allowing us to quickly capture information about college students’ use of context when identifying PTSD and MDD, before employing open-ended responses that must be coded before analyses. This study could have benefitted from a larger sample size, but again, was only a pilot study, and a larger number of subjects were recruited in the main study.
Appendix D

Clinician Open-Ended Disorder Identification Responses

This section describes all of the open-ended diagnoses made by clinicians and clinical trainees. Here, we report diagnoses made in Experiment 1 and Experiment 2 for PTSD, MDD, and distressed behaviors.

In Experiment 1, 136 people saw six vignettes each; in Experiment 2, 71 people judged three vignettes each in the open-ended diagnosis task. Thus, clinicians and clinical trainees judged a total of 1029 vignettes. For all 1029 vignettes, people offered 1427 responses; 512 for PTSD behaviors, 529 for MDD behaviors, and 386 for distressed behaviors. People were sometimes willing to make more than one response for a single vignette; on average, there were 1.39 responses each. Ninety-eight percent of participants’ responses were a DSM disorder diagnosis or an indication that no diagnosis was necessary (“none”). Two percent were not.

**DSM Disorder Identifications**

Across vignettes, clinicians and trainees offered 23 distinct DSM-IV-TR (APA, 2000) or DSM-5 (APA, 2013) disorder diagnoses: PTSD, MDD/Depression, Adjustment Disorder, Acute Stress Disorder, Anxiety Disorder, Generalized Anxiety Disorder, Panic Disorder, Agoraphobia, Dysthymia, Personality Disorder, Bipolar Disorder, Unspecified Trauma and Stressor Related Disorder, Specific Phobia, Eating Disorder, Binge Eating Disorder, Borderline Personality Disorder, Mood Disorder, Obsessive Compulsive Disorder, Social Anxiety Disorder, Insomnia Disorder, Depersonalization / Derealization Disorder, Dissociative Disorder, and Schizoaffective Disorder. Please see Table C1 for details by behavioral vignette.

Participants made the “appropriate” diagnosis for 82% of vignettes. That is, they indicated the disorder we described (as carefully as possible) for the disordered behavioral
symptoms (PTSD or MDD), or they indicated “none” for the distressed vignette. Again, each vignette fulfilled all of the diagnostic criteria for either a PTSD diagnosis, a MDD diagnosis, or no disorder at all. Overall, they indicated PTSD for 74% of PTSD behavioral vignettes, depression or MDD for 87% of MDD behavioral vignettes, and indicated “none” for 85% of distressed behavioral vignettes. Also, participants diagnosed 26% of distressed vignettes with a DSM disorder (at times in addition to a no-diagnosis indication), suggesting a potential tendency to over-diagnose (with respect to DSM guidelines).

Multiple DSM disorder identifications for single vignettes. Clinicians and trainees indicated two or three DSM disorder identifications for 32% of vignettes, 37% for PTSD vignettes, 47% for MDD vignettes, and 11% for distressed vignettes. Some participants may have been intending to make comorbidity diagnoses. Indeed, PTSD and MDD are often comorbid (DSM-5; APA, 2013; Hasin, Goodwin, Stinson, & Grant, 2005). On the other hand, people may have meant to indicate multiple alternative diagnoses.

“No Diagnosis” Responses

Participants infrequently indicated that no disorder was present (“none”) for PTSD behaviors (3%), and MDD behaviors (<1%). The vast majority of “none” responses were made “appropriately” for non-disordered distressed behaviors (85%). Thus, clinicians and clinical trainees appeared to carefully assess the behavioral vignettes for whether a disorder was present at all.

Non-DSM Disorder Identifications

The remaining 2% of responses were neither DSM disorders nor “none” responses. People made a total of 25 such responses; 10 for PTSD behaviors, 10 for MDD behaviors, and 5 for distressed behaviors. Interestingly, five were trauma-related responses likely motivated by
traumatic event context: specific fear of military personnel (1), secondary trauma (1), victim of crime (2), and Exposure to Disaster, War, or other Hostilities (1). Seven responses simply indicated bereavement (6), or grief (1). An additional seven responses resembled a proposed DSM disorder, and thus may still be viewed as a diagnosis: complicated bereavement disorder (1), complicated grief (1), other specified obsessive-compulsive and related disorder – compulsive shopping (1), impulse control disorder (2), intermittent explosive disorder (2). Five responses resembled a DSM disorder: anxiety reaction (1), adjustment-like disorder (1), other specified depressive disorder (insufficient symptoms) (1), unspecified depressive disorder (1), and depressive reaction (1). Finally, one participant diagnosed a medical condition, hypothyroidism, for MDD behaviors.

The Influence of Context on Adjustment Disorder Diagnoses: Secondary Analyses

As Adjustment Disorder (please see Footnote 1) was the only diagnosis besides those examined in the main paper that was made by at least ten participants per behavior type, we ran chi square tests to examine the effect of event context (traumatic event, everyday event) and reaction context (intense reaction, mild reaction) for all three behavior types (PTSD, MDD, distressed) on the number of Adjustment Disorder diagnoses across Experiments 1 and 2. Further, participants judged the Adjustment Disorder diagnoses examined below to be the most likely diagnosis, or the only one, for a given vignette.

Chi square tests revealed that clinicians and clinical trainees indicated Adjustment Disorder as the most likely diagnosis for PTSD behaviors more often in the context of an intense reaction (18 out of 25 diagnoses) than in the context of a mild reaction (7 out of 25 diagnoses; \(X^2[1, N = 343] = 5.29, p = .021\)). Event context did not affect Adjustment Disorder diagnoses \((X^2[1, N = 343] = .05, p = .824)\). In addition to the PTSD diagnoses that were made for these
PTSD behaviors, the presence of an intense reaction before PTSD behaviors may have made it seem more likely to clinicians and clinical trainees that the person in the vignette was responding and adapting immediately to the event, thereby increasing Adjustment Disorder diagnoses in intense reaction context.

Participants’ diagnoses of Adjustment Disorder for MDD and distressed behaviors were not influenced by event or reaction contexts (all $ps \geq .435$). Out of all 343 MDD behavioral vignettes, Adjustment Disorder was rated as the most likely diagnosis 30 times (9%). For all 343 distressed behavioral vignettes, Adjustment Disorder was rated most likely 18 times (5%).
### Table C1

*All Open-Ended Diagnosis Responses by Clinicians and Clinical Trainees in Experiment 1 and Experiment 2*

<table>
<thead>
<tr>
<th>Open-Ended Diagnosis Responses</th>
<th>Behavior Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PTSD</td>
</tr>
<tr>
<td>PTSD</td>
<td>254</td>
</tr>
<tr>
<td>Depression (including MDD)</td>
<td>68</td>
</tr>
<tr>
<td>&quot;None&quot;</td>
<td>11</td>
</tr>
<tr>
<td>Adjustment Disorder</td>
<td>36</td>
</tr>
<tr>
<td>Acute Stress Disorder</td>
<td>26</td>
</tr>
<tr>
<td>Anxiety Disorder</td>
<td>26</td>
</tr>
<tr>
<td>Generalized Anxiety Disorder</td>
<td>22</td>
</tr>
<tr>
<td>Panic Disorder</td>
<td>8</td>
</tr>
<tr>
<td>Agoraphobia</td>
<td>13</td>
</tr>
<tr>
<td>Dysthymia</td>
<td>1</td>
</tr>
<tr>
<td>Personality Disorder</td>
<td>4</td>
</tr>
<tr>
<td>Bipolar Disorder</td>
<td>9</td>
</tr>
<tr>
<td>Unspecified Trauma and Stressor Related disorder</td>
<td>5</td>
</tr>
<tr>
<td>Specific Phobia</td>
<td>7</td>
</tr>
<tr>
<td>Binge Eating Disorder</td>
<td>6</td>
</tr>
<tr>
<td>Eating Disorder</td>
<td>4</td>
</tr>
<tr>
<td>Disorder</td>
<td>DSM-IV-TR</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Borderline Personality Disorder</td>
<td>2</td>
</tr>
<tr>
<td>Mood Disorder</td>
<td>2</td>
</tr>
<tr>
<td>Obsessive Compulsive Disorder</td>
<td>3</td>
</tr>
<tr>
<td>Social Anxiety Disorder</td>
<td>2</td>
</tr>
<tr>
<td>Insomnia Disorder</td>
<td>1</td>
</tr>
<tr>
<td>Depersonalization / Derealization Disorder</td>
<td>1</td>
</tr>
<tr>
<td>Dissociative Disorder</td>
<td>1</td>
</tr>
<tr>
<td>Schizoaffective Disorder</td>
<td>1</td>
</tr>
<tr>
<td>Non DSM-IV-TR or DSM-5 Disorder Responses</td>
<td>9</td>
</tr>
</tbody>
</table>

| Total Responses                              | 512       | 529   | 386                   |
Appendix E
Supplementary Experiment: Clinicians’ Diagnoses and Judgments of Treatment Need, Distress, Functional Impairment, and Psychological Abnormality

We conducted a supplementary experiment to explore how it is that traumatic event context made PTSD behaviors seem less psychologically abnormal, but more in need of a diagnosis in Experiment 1 of the clinician manuscript. In clinical psychology, there has been much philosophical discussion about what factors constitute a diagnosable mental disorder, and what considerations point towards the need for a diagnosis (Bennett, 2003; Bennett, 2011; Kessler et al., 2005; Kring, Johnson, Davison, & Neale, 2010; Lilienfeld & Marino, 1995). This philosophical discussion has focused on factors that have been argued to align with diagnosis need. Treatment need refers to the need for repair, alleviation, intervention, or care. Distress concerns an individual’s internal suffering, such as guilt or anxiety. Functional impairment means disability in a major area of life, to the same extent that one might experience with a physical injury, such as inability to work or fulfill key familial roles. Finally, psychological abnormality (which clinicians in Experiment 1 were asked to judge) is characterized by distorted psychological processes. Each of these four factors has been argued by different camps to be critical in identifying behaviors that are truly mentally disordered, and that therefore merit a mental disorder diagnosis (Bennett, 2003; Bennett, 2011; Kessler et al., 2005; Kring, Johnson, Davison, & Neale, 2010; Lilienfeld & Marino, 1995). Using these ideas as a starting point, we wanted to better understand which factors actually guide clinicians’ diagnosis decisions, if not psychological abnormality.

As we saw in Experiment 1, considerations of psychological abnormality did not seem to influence PTSD diagnoses. In the current exploratory study, we again measured a new set of
clinicians’ and clinical trainees’ assessments of psychological abnormality and the need for a PTSD diagnosis (using a newly modified dependent variable; see below) in addition to three other assessments that might contribute to diagnosis decisions: treatment need, distress, and disability and/or functional impairment (henceforth referred to simply as functional impairment). We were primarily interested in whether judgments of treatment need, distress, functional impairment, and psychological abnormality can explain the variance in people’s assessments of PTSD diagnosis. These data could illuminate why causal context makes PTSD behaviors seem more in need of a diagnosis even while seeming less psychologically abnormal. Simultaneously, we also examined whether the presence of traumatic event context affects these assessments in the same way that clinicians’ and clinical trainees’ PTSD diagnoses were affected by traumatic event context in Experiment 1.

Finally, we expanded upon the Experiment 1 findings by exploring yet another issue: Whether clinicians’ judgments are influenced only by the type of traumatic event specified in the DSM-5 (APA, 2013) as meeting the A1 criterion (e.g., threatened death, serious injury). To do this, we examined the influence of pre-rated traumatic events both meeting and not meeting the A1 criterion, relative to the influence of everyday events, on clinical judgments. The results of Experiment 1 suggested that clinicians seem to make diagnoses in a manner consistent with the DSM-5 (APA, 2013), and this manipulation would allow us to further explore that possibility.

Thus, we asked clinicians and clinical trainees to make judgments of a hypothetical person’s behavioral symptoms in the context of his or her recent experiences, following the design of Experiment 1 and using the materials of Experiment 1, except for the following alterations.
First, we employed different dependent variables in the current exploratory study. The open-ended diagnosis task from Experiment 1 was replaced with a measure asking participants to judge the need for a PTSD diagnosis (i.e., on a Likert scale). Standardized screening instruments often incorporate rating scales, so these kinds of measurements are, in fact, a realistic part of clinical practice (Lang & Stein, 2005). Also, a closed-ended diagnosis judgment measure, in which participants are asked to merely indicate whether a PTSD diagnosis is merited, rather than consider and articulate one or more diagnoses, shortened the task relative to the open-ended diagnosis judgment measure. For the “Subsequent Behaviors” in each vignette, we asked people to make nine total assessments: (1) likelihood of a PTSD diagnosis, (2) likelihood of a MDD diagnosis, (3) likelihood of any diagnosis overall (not explicitly excluding PTSD and/or MDD), (4) treatment need for PTSD, and (5) treatment need for MDD, (6) perceived distress, (7) perceived functional impairment, (8) psychological abnormality, and (9) difficulty understanding (please see Table D1 for exact dependent variable text).

Second, we manipulated two types of traumatic events: Traumatic events as defined by the A1 criterion (e.g., threatened death) in the *DSM-IV-TR* (APA, 2000) and *DSM-5* (APA, 2013), versus non-A1 criterion events that have nevertheless been shown to lead to PTSD behavioral symptoms (e.g., divorce; Helzer, Robins, & McEvoy, 1987; burglary without confrontation with the burglar, and unemployment; Mol et al., 2005).

Third, because reaction information did not influence people’s judgments in Experiment 1, we did not include reaction information in the “Recent Past” paragraph and included only

---

6 In Experiment 1, nine clinicians began but did not finish the survey, perhaps due to the overall length of the task, and several complained about its length. Thus, we felt that task length was a potential retention concern.
event information. Instead, we further explored clinicians’ processing of reaction information in Experiment 2.

Fourth, we asked people to reason about three vignettes instead of six to help lower the total experiment time, given that they were to give judgments for a greater number of dependent measures. One vignette each described PTSD behaviors, MDD behaviors, and distressed behaviors.

**Expected Replications**

We expected to replicate the following Experiment 1 and 2 findings: PTSD behaviors would be rated as more likely to need a PTSD diagnosis in traumatic event context compared to everyday event context, MDD diagnosis likelihood judgments for MDD behaviors would not be affected by event context, and distressed behaviors would either seem more in need of a *DSM* disorder diagnosis in traumatic event context (Experiment 1) or no more or less in need of a *DSM* diagnosis in context (Experiment 2). We also predicted that, as in Experiment 1, PTSD and MDD behaviors would seem less psychologically abnormal and less difficult to understand in traumatic event context, compared to everyday event context, whereas distressed behaviors would seem more psychologically abnormal and more difficult to understand in traumatic event context, compared to everyday event context.

We asked three key questions in this exploratory study (Questions A1, A2, and A3).

**Question A1: What factors predict PTSD diagnosis decisions, if not psychological abnormality judgments?**

**Treatment need.** We aimed to test whether PTSD treatment need judgments predict a substantial part of the variance in PTSD diagnosis judgments. Perceived PTSD treatment need might indeed predict PTSD diagnosis judgments if clinicians identify cases of PTSD in part by
pinpointing people with symptoms that are particularly in need of alleviation, repair, treatment or care for PTSD specifically. On the other hand, regression analyses might not provide any evidence that perceived treatment need predicts PTSD diagnosis need judgments. That is, clinicians may be able to identify PTSD in a hypothetical person without necessarily believing that person needs to be treated.

**Distress.** We also used regression analyses to test whether perceived distress predicts PTSD diagnosis need judgments. This may be the case if clinicians believe that PTSD ought to be diagnosed especially when a person appears to be experiencing intense internal distress. If so, then clinicians might consider how distressing a set of behaviors are when deciding whether those behaviors merit a PTSD diagnosis. Alternatively, regression analyses might show that perceived distress does not predict PTSD diagnosis need judgments. For example, clinicians may feel that distress is ubiquitous across mental disorders, and is not an especially strong indicator that PTSD is present. Indeed, distress itself is specified as a diagnostic criterion in some DSM mental disorders (e.g., MDD), but not PTSD. Thus, clinicians might diagnose PTSD regardless of how much distress they feel the hypothetical person is experiencing.

**Functional impairment.** We also tested whether perceived functional impairment explains any of the variance in diagnosis need judgments. One reason this might be the case is that clinicians are likely to know of specific functional impairments that sometimes accompany PTSD behaviors. For example, agoraphobia is known to sometimes co-occur with PTSD, which can limit people from leaving the house (Maes, Mylle, Delmeire, Altamura, 2000). Thus, functional impairment is a relevant consideration that clinicians might wish to pinpoint when diagnosing PTSD. On the other hand, clinicians might believe that functional impairment is not necessarily a good indicator of PTSD; for example, someone with PTSD might be able to
function with mild to moderate behavioral modifications to avoid trauma reminders. So, functional impairment considerations may be more or less irrelevant to clinicians who are considering a PTSD diagnosis.

**Psychological abnormality.** Given the Experiment 1 results, we expected to find that psychological abnormality judgments do not positively correlate with diagnosis likelihood judgments (please see an extensive discussion of these findings in the General Discussion section of the clinician manuscript).

In addition, we examined Question A1 for MDD, to investigate whether the expected findings for PTSD behaviors also extended to another disorder. One possibility is that treatment need, distress, functional impairment, and/or psychological abnormality judgments will predict judgments of MDD diagnosis likelihood for MDD behaviors, or not predict them, in the same way as for PTSD behaviors. These specific considerations, whichever they turn out to be, may map uniformly to clinical considerations of all disorders. That is, regardless of which factors are thought to align with diagnosis need according to philosophical discussion (e.g., Bennett, 2003) only a specific subset might align with diagnosis need judgments for both disorders tested in this study, because these are some underlying factors that guide clinicians’ diagnoses regardless of the specific symptoms present in a given case. For example, treatment need may be a particularly important consideration when making PTSD diagnoses and MDD diagnoses, because clinicians and clinical trainees reason that both cases are sufficiently severe to merit treatment.

An alternative is that treatment need, distress, functional impairment, and/or psychological abnormality judgments will predict judgments of MDD diagnosis likelihood in a configuration that does not match the factors that predict judgments of PTSD likelihood for PTSD behaviors. That is, there may not be disorder-general considerations that predict diagnosis
need judgments because clinical professionals feel different disorders need to be diagnoses for different reasons. For example, treatment need may be a particularly important consideration when making PTSD diagnoses because clinicians may reason that PTSD symptoms are unlikely to remit on their own, but treatment need may not predict MDD diagnoses, which clinicians may reason some people overcome on their own by developing lifestyle changes, seeking social support, or simply over time.

**Question A2: Does non-A1 traumatic event context increase PTSD diagnoses to the same extent that A1 traumatic event context increases them, relative to everyday event context?**

One possibility is that participants will not diagnose PTSD for PTSD behaviors that occur in the presence of a non-A1 traumatic event (e.g., one that is not life-threatening, such as divorce or unemployment) because they have internalized this longstanding diagnostic criterion for PTSD, and will judge that the non-A1 traumatic events do not qualify the hypothetical person for a PTSD diagnosis. The A1 criterion did not change dramatically from the *DSM-IV-TR* (APA, 2000) to the *DSM-5* (APA, 2013); thus, clinicians and clinical trainees, all of whom should be familiar with at least one version, may not diagnose PTSD unless the official A1 criterion (i.e., involving threatened death) for a qualifying traumatic event is clearly fulfilled.

A plausible alternative prediction is that participants will diagnose PTSD for PTSD behaviors in the context of non-A1 traumatic events because they will interpret these events as sufficiently severe and negative to trigger the onset of PTSD symptoms (Horwitz & Wakefield, 2007). Such a finding would suggest a relative departure from *DSM*-prescribed diagnosis.

To examine whether any effects obtained for PTSD behaviors reflected a disorder-general tendency or a behavior-general tendency, we also examined whether the two types of traumatic event context influenced diagnosis likelihood judgments for MDD behaviors and
distressed behaviors. However, we expected to find that diagnosis judgments for MDD behaviors would not be affected by event context, as in Experiments 1 and 2. We expected to find either that distressed behaviors would seem more in need of a DSM disorder diagnosis in traumatic event context, as in Experiment 1, or would be unaffected by event context, as in Experiment 2.

**Question A3:** Are assessments of treatment need, distress, and functional impairment for PTSD behaviors increased or decreased by traumatic event context, compared to everyday event context?

**Treatment need.** One hypothesis is that traumatic life event context for PTSD behaviors will lead to higher assessments of PTSD treatment need. Participants may feel, for example, that when it is a traumatic event that yields PTSD behaviors, those behaviors are more PTSD-like and disorder-like than if an everyday event yielded the same behaviors. An alternative is that traumatic life event context will lead people to view PTSD behaviors as less in need of treatment. Weine, Kim, & Lincoln (2015) found that traumatic event context decreased treatment need judgments for a hypothetical patient who met DSM-IV-TR (APA, 2000) criteria for an alcohol use disorder, in the opinions of lay people. Kim & LoSavio (2009) similarly found that a hypothetical person with randomly sampled disordered behaviors from the DSM-IV-TR (APA, 2000) was rated as less in need of treatment in the opinions of lay people, when they were given an externally-controlled life event explanation for that person’s behaviors (compared to an internally-controlled explanation). It is unknown, however, whether clinicians may be similarly influenced by traumatic event context, viewing PTSD behaviors as less in need of treatment for PTSD.

**Distress.** For judgments of distress, one plausible hypothesis is that traumatic event context will increase judgments that PTSD behaviors are distressed, just as traumatic context
increased participants’ PTSD diagnoses in Experiment 1. This would suggest that people find someone who recently experienced a traumatic event, followed by PTSD behaviors, to be experiencing more distress than someone who recently experienced an everyday event, followed by PTSD behaviors, perhaps because this person is dealing with more negative issues overall. An alternative hypothesis is that traumatic event context will decrease judgments that PTSD behaviors are distressed just as it decreased psychological abnormality judgments in Experiment 1. Clinicians and clinical trainees might perceive PTSD behaviors, viewed in proportionate context to a traumatic event, are in comparison not unreasonable.

**Functional impairment.** Finally, for judgments of functional impairment, one hypothesis is that traumatic event context will increase judgments that someone with PTSD behaviors is experiencing functional impairment. PTSD behavioral symptoms themselves impair daily functioning. However, a traumatic event plus PTSD behaviors might to lead participants to perceive greater functional impairment than an everyday event plus PTSD behaviors. One reason for this is clinicians may infer that the traumatic event itself (e.g., relocation, injury or death of a loved one) leads that person to experience more disability in daily life, compared to the relative lack of downstream effects stemming from an everyday event itself (e.g., reading a story in the newspaper; witnessing a fender bender). Or, they may feel that someone whose PTSD behaviors occurred after an everyday event may have developed those PTSD behaviors because they were psychologically or behaviorally unequipped to deal with even everyday events well, and thus are more likely to be experiencing functional impairment.

An alternative hypothesis for judgments of functional impairment is that traumatic event context will decrease judgments that someone with PTSD behaviors is functionally impaired. Participants’ perceptions of functional impairment may be decreased by traumatic event context
via the understanding-normality effect (Ahn et al., 2003). More specifically, clinicians and clinical trainees may feel that someone’s PTSD behaviors are an understandable response to a traumatic event, and do not suggest a lack of functional impairment per se.

Methods

Participants.

Clinicians. Thirty-one licensed practicing clinicians (22 female) participated in response to a mailed postcard invitation to an online study. As a token of thanks for participation, clinicians who completed the study were sent a $25 gift certificate to an online retailer via an email address they voluntarily provided. A random sample of 200 clinicians from 47 states in the U.S. was taken, in the same way as Experiment 1. Out of the 200 clinicians who were mailed postcards, eight clinicians’ cards were returned because the clinician had moved. An additional two clinicians who participated were excluded from the final count of 31 after failing to complete a number of the key dependent measures. This yielded a 17.2% response rate, which is comparable to past work using a similar recruitment method to sample a clinical population (Rottman, Ahn, Sanislow, & Kim, 2009).

Clinicians reported a mean age of 43 years (range: 26-71), and self-identified as White (90%), Black (3%), and American Indian (3%; the remaining participants declined to report race). In addition, 7% self-identified as Hispanic. Most clinicians held a private practice (74%) but also practiced in other settings (10% hospital, 7% community clinic, 3% specialized center for autism, 3% university mental hygiene clinic, 3% psychology department clinic). Clinicians had an average of 18.3 years of clinical experience (range 2-40 years) and held a range of primary clinical orientations (48% Cognitive / Cognitive-Behavioral, 19% Psychoanalytic, 13% Humanistic, 10% Eclectic, 7% Family Systems, and 3% Behavioral).
Clinical trainees. Twenty-five graduate students (20 female) were also recruited in the same way as Experiment 1. Graduate students who completed the study were entered into one of several raffles for a $25 gift card to an online retailer, using an email address they voluntarily provided. Graduate students reported a mean age of 27 years (range: 23 - 34), and self-identified as White (88%), Asian (8%), and multiple races (4%). No participants self-identified as Hispanic. They had an average of 2.0 years of clinical experience (range 0-6 years) and provided clinical services in a variety of settings (44% psychology department clinics, 20% hospitals, 8% community clinics, 4% private practice, 12% other, 12% no service experience). Graduate students’ primary clinical orientation was most frequently Cognitive / Cognitive-Behavioral (76%), followed by Behavioral (8%) and others (4% Humanistic, 4% Family Systems, 4% Eclectic, 4% declined to report).

Non-A1 traumatic event norming. Three non-A1 criterion traumatic event vignettes were created to complement three vignettes from Experiment 1 (“Martin,” “Diana,” and “Justin”). Each non-A1 traumatic event was written to describe a hypothetical person from the original vignettes, this time experiencing a non-A1 criterion traumatic event, which has been shown to lead to PTSD symptoms. “Martin’s” non-A1 traumatic event vignette describes divorce (Helzer et al., 1987), “Diana’s” non-A1 traumatic event describes burglary without confrontation with the burglar (Mol et al., 2005), and “Justin’s” non-A1 traumatic event vignette describes sudden unemployment (Mol et al., 2005). Please see Table D2 for all three non-A1 traumatic event vignette texts. The traumatic event text, everyday event text, and behavioral symptoms were preserved from the original Experiment 1 vignettes for “Martin” “Diana” and “Justin” (please see Appendix A). Thus, the key between-subject factor is Event Type (Non-A1 traumatic event, A1 traumatic event, Everyday event).
Sixteen people (12 female) with formal training in clinical psychology (12 psychiatry residents, 3 postdoctoral fellows in clinical psychology, and 1 psychiatric nursing intern) who did not participate in the main study independently rated each A1 traumatic event, each non-A1 traumatic event, and each everyday event for how traumatic it seemed, using the same method as the Experiment 1 manipulation check. Because these materials were developed in separate phases, eight participants rated the vignettes about “Martin” and eight participants rated the vignettes about “Diana” and “Justin.” Paired samples t-tests confirmed that participants did not perceive the three A1 traumatic events to be either more or less traumatic than the non-A1 traumatic events (Martin: \( t[7] = .21, p = .844 \); Diana: \( t[7] = 0.00, p = 1.000 \); Justin \( t[7] = 1.80, p = .116 \)). Also, participants also found all three non-A1 traumatic events to be more traumatic than the everyday events (Martin: \( t[7] = 6.42, p < .001 \); Diana: \( t[7] = 8.79, p < .001 \); Justin \( t[7] = 2.54, p = .039 \)). Just as we found in the Experiment 1 manipulation check, participants also found all three A1 traumatic events to be more traumatic than the everyday events (Martin: \( t[7] = 5.17, p = .001 \); Diana: \( t[7] = 7.34, p < .001 \); Justin \( t[7] = 6.54, p < .001 \).

**Procedure.** In the main study, all vignettes and questions were presented in counterbalanced order, such that people assessed one vignette for all nine dependent variables in randomized order, followed by the second vignette for all nine dependent variables, until all questions had been answered for all vignettes. Vignettes themselves were also always presented in randomized order, and there was always one vignette and one question per page.

**Results and Discussion**

**Multiple regression analyses: Examining factors that predict PTSD diagnoses.** To address Question A1, multiple linear regression analyses were conducted to examine whether judgments of treatment need, distress, functional impairment, and psychological abnormality
predict whether or not PTSD behaviors and MDD are judged to merit a disorder-specific diagnosis.

**Primary analyses: PTSD behaviors.** For PTSD behaviors, PTSD treatment need was positively correlated with ratings of PTSD likelihood ($r[53] = .662, p < .001$). Perceived distress was also positively correlated with ratings of PTSD likelihood ($r[53] = .254, p = .031$), as was perceived functional impairment ($r[53] = .260, p = .027$). In contrast, psychological abnormality ratings were uncorrelated with ratings of PTSD likelihood ($r[53] = -.048, p = .364$), as hypothesized. Please see Table D3 for a correlation matrix.

For PTSD behaviors, a multiple regression analysis examining all four of these factors simultaneously was performed using the enter method on PTSD likelihood judgments. The resulting model explained a significant amount of the variance in PTSD likelihood ratings ($R^2 = .465, F[4,50] = 10.89, p < .001$; please see Table D4). PTSD treatment need significantly predicted PTSD likelihood ratings for PTSD behaviors ($\beta = .609, t[50] = 5.88, p < .001$), whereas distress, functional impairment, and psychological abnormality did not (all $ps \geq .147$). This model emerged regardless of the order in which the factors were entered. These results demonstrate, first, that considerations of PTSD diagnosis likelihood seem to align substantially with judgments of treatment need, but not with judgments of psychological abnormality. In addition, judgments of distress and functional impairment correlate with PTSD diagnosis likelihood judgments, but do not predict PTSD diagnosis likelihood judgments independently of judgments of treatment need. This adds distress and functional impairment to the list of judgments that do not align with PTSD diagnosis, along with psychological abnormality, as our Experiment 1 data suggested. However, one notable difference between psychological abnormality and distress and functional impairment is that psychological abnormality did not
correlate with PTSD diagnosis judgments in this study, whereas the other two factors did. One way of interpreting this is distress and functional impairment may be considerations clinicians and clinical trainees take into account when making PTSD diagnoses, but ultimately do not guide their decisions to the extent that considerations of treatment need do so.

**Secondary analyses: MDD behaviors.** We also conducted the same analyses for MDD behaviors, to understand whether the above findings for PTSD might be extended to other disorders. For MDD behaviors, MDD diagnosis likelihood was significantly correlated with MDD treatment need ($r[53] = .642, p < .001$), perceived distress ($r[53] = .548, p < .001$), perceived functional impairment ($r[53] = .473, p < .001$), and psychological abnormality judgments ($r[53] = .325, p = .008$). Please see Table D5 for the full correlation matrix.

For MDD behaviors, a regression analysis examining all four of these factors simultaneously was performed using the enter method on MDD likelihood judgments. The model that emerged explained a significant amount of the variance in MDD likelihood ratings ($R^2 = .480, F[4,50] = 11.56, p < .001$; please see Table D6). MDD treatment need significantly predicted MDD likelihood judgments for MDD behaviors ($\beta = .451, t[50] = 3.62, p = .001$), whereas distress, functional impairment, and psychological abnormality did not (all $ps > .297$). Again, this model emerged regardless of which factor was entered first. These results demonstrate that considerations of MDD diagnosis likelihood for MDD behaviors are best predicted by MDD treatment need considerations. Distress, functional impairment, and psychological abnormality predicted MDD diagnosis likelihood separately, but not over and above the predictive power of MDD treatment need. Thus, MDD treatment need is a highly relevant dimension when considering whether certain behaviors are indicative of MDD (or conceivably vice-versa).
Indeed, the same clinical consideration, treatment need, guided diagnoses of PTSD and MDD, over and above the influence of the other factors we examined (Question A1). Thus, these data suggest that treatment need may guide clinicians when making disorder diagnoses in general. If so, we might expect clinicians to diagnose disorders with fewer available treatment options less frequently than disorders with many well-known treatment options. More research is needed to examine this possibility, and to see whether treatment need guides disorder-specific diagnosis judgments for other mental disorders in the DSM, so that we can better understand whether this is truly a disorder-nonspecific effect.

Analyses of variance: Examining the influence of A1 and non-A1 traumatic event context on PTSD, MDD, and non-disordered diagnosis likelihood judgments. To examine Question A2, whether the assessments of PTSD diagnosis likelihood for PTSD behaviors were affected by A1 traumatic event context versus non-A1 traumatic event context, we ran three separate ANOVAs. Specifically, we ran three separate 3 (Event Context: A1 traumatic event, non-A1 traumatic event, everyday event) x 2 (Expertise: Clinician, Clinical trainee) ANOVAs on PTSD diagnosis likelihood judgments for PTSD behaviors, MDD diagnosis likelihood judgments for MDD behaviors, and any DSM diagnosis likelihood judgments for distressed behaviors. To see whether traumatic event context (A1 and non-A1) affected judgments differently from the control everyday event behaviors, any main effects of event context, or interactions involving event context, were further investigated with two Bonferroni-corrected independent samples t-tests. One t-test compared A1 traumatic event context and everyday event context and the other t-test compared non-A1 traumatic event context and everyday event context. The unexamined judgments (e.g., MDD diagnosis likelihood for PTSD behaviors) simply served to reduce potential demand characteristics by ensuring that participants were not
only asked the diagnosis and treatment need questions of interest for each behavior type (e.g., only MDD diagnosis likelihood for MDD behaviors).

**Primary analyses: PTSD diagnosis likelihood judgments for PTSD behaviors.** An ANOVA on PTSD diagnosis likelihood judgments for PTSD behaviors revealed a main effect of event context qualified by an interaction of event context and expertise ($F[2,49] = 3.81, p = .029$), please see Figure D1. There was no main effect of expertise ($p = .132$). To examine the interaction, Bonferroni-corrected t-tests were performed on clinicians’ and clinical trainees’ judgments separately. For clinicians, there was no difference between PTSD diagnosis judgments in either type of traumatic event context compared to everyday event context (A1 traumatic event context and everyday event context: $t[20] = -.10, p = .921$; non-A1 traumatic event context and everyday event context: $t[17] = -1.72, p = .104$). However, clinical trainees judged PTSD behaviors in A1 traumatic event context to be more likely to have PTSD ($M = 7.71, SD = .76$) compared to PTSD behaviors in everyday event context ($M = 4.70, SD = 2.06; t[15] = -3.68, p = .002$). Clinical trainees did not judge PTSD behaviors in non-A1 traumatic event context to be any more likely to have PTSD than in everyday event context ($t[16] = -.87, p = .016$). Thus, clinical trainees judged PTSD behaviors to be more likely to have PTSD in A1 traumatic event context than in everyday event context (Question A3), but clinicians were not differently influenced by context when judging the diagnosis likelihood of PTSD behaviors in this study.

We did not find such an interaction of event context and expertise in Experiment 1; instead, both clinicians and clinical trainees found PTSD behaviors to be more likely to merit a PTSD diagnosis in A1 traumatic event context compared to everyday event context. Possibly, the closed-ended measure we used in Appendix E led clinicians to pay close attention to the
behavioral symptoms, perhaps to count how many there were, and thus to be relatively less influenced by event context than they were in Experiment 1. Clinical trainees may have been adhering more strictly to the DSM-5 criteria (APA, 2013) than clinicians were, leading them to rate a PTSD diagnosis more likely for PTSD behaviors in A1 traumatic event context (DSM-IV-TR, APA, 2000; DSM-5, APA, 2013) but not in non-A1 traumatic event context. Again, please see Figure D1.

Overall, the non-A1 traumatic events were seen as just as likely to be indicative of PTSD as traumatic events (Question A2). Indeed, an independent-samples t-test revealed no significant difference between PTSD diagnosis likelihood judgments in non-A1 traumatic event context and A1 traumatic event context (t[32] = .467, p = .643). Thus, people with PTSD behaviors that occur after a non-A1 traumatic event may be just as likely to receive a PTSD diagnosis in clinical practice even though this type of event does not fit the DSM-5 (APA, 2013) diagnostic criteria for PTSD. At least some clinicians and clinical trainees may be aware of the literature on non-A1 traumatic events, from which all of our non-A1 traumatic event vignettes were drawn. If so, they may have diagnosed PTSD in line with this literature. On the other hand, PTSD behaviors in non-A1 traumatic event context were not seen as any more likely to be PTSD than PTSD behaviors in everyday event context. So, it may yet be the case that people with PTSD behavioral symptoms that occur after a non-A1 traumatic event are less likely to receive a PTSD diagnosis in clinical practice compared to those who have PTSD behavioral symptoms in A1 traumatic event context. Future work is needed to investigate these possibilities in a more ecologically valid setting.

**Secondary analyses: MDD diagnosis likelihood judgments for MDD behaviors.** The ANOVA on MDD diagnosis likelihood judgments for MDD behaviors revealed no significant
main effects or an interaction (all \( ps > .303 \)). The mean MDD diagnosis likelihood rating was 7.80 (\( SD = 1.17 \)) on a scale of 1-9 (where 1 = MDD diagnosis very unlikely, and 9 = MDD diagnosis very likely). Indeed, a one sample t-test demonstrated that these judgments were not only above the midpoint of “5,” but also significantly above a “7” on this scale (\( t[55] = 5.16, p < .001 \)). Thus, clinicians and clinical trainees generally judged MDD to be a likely diagnosis for MDD behaviors. This finding aligns with our Experiment 1 result that participants were not influenced by context when diagnosing MDD or depression for MDD behaviors. Also, this finding demonstrates that the increase of PTSD diagnosis likelihood judgments made by clinical trainees in A1 traumatic event context did not extend to judgments for MDD behaviors, suggesting that the effect found for PTSD behaviors was disorder-specific.

**Secondary analyses: Any DSM diagnosis likelihood judgments for distressed behaviors.** The ANOVA on any DSM diagnosis likelihood judgments for distressed behaviors revealed no significant effects (all \( ps > .492 \)). The mean was a 2.47 (\( SD = 1.59 \)) on a scale of 1-9 (where 1 = any diagnosis very unlikely, and 9 = any diagnosis very likely). Indeed, a one sample t-test demonstrated that these ratings were below not only the midpoint, which was “5,” but were also significantly below a “3” on this scale (\( t[54] = -2.47, p = .017 \)). Thus, clinicians and clinical trainees generally judged distressed behaviors to be unlikely to need any DSM diagnosis. This aligns with our Experiment 2 finding that participants were uninfluenced by context when making diagnoses for distressed behaviors, but not by the Experiment 1 finding that event context increased diagnoses for these behaviors. However, participants diagnosed fewer distressed vignettes in Experiment 2 (71) and the current study (56) than they did in Experiment 1 (343). Quite possibly, the effect of event context on diagnoses for distressed behaviors was only apparent among the relatively much larger amount of judgments.
Analyses of variance: The influence of A1 and non-A1 traumatic event context on additional judgments for PTSD behaviors. To examine Question A3, whether assessments other than diagnosis likelihood were affected by A1 traumatic event context versus non-A1 traumatic event context, we separate 3 (Event Context: A1 traumatic event, non-A1 traumatic event, everyday event) x 2 (Expertise: Clinician, Clinical trainee) ANOVAs on the following assessments of theoretical interest: For PTSD behaviors, we analyzed judgments of the likelihood of a PTSD diagnosis and treatment need for PTSD, judgments of perceived distress, perceived functional impairment, psychological abnormality, and difficulty understanding. Main effects of event context, or interactions involving event context, were further investigated with Bonferroni-corrected independent samples t-tests. Again, the unexamined judgments (e.g., MDD treatment need likelihood for PTSD behaviors) were present to help reduce demand characteristics.

Primary analyses: PTSD treatment need judgments for PTSD behaviors. The ANOVA on PTSD treatment need judgments for PTSD behaviors revealed a main effect of event context qualified by an interaction of event context and expertise (F[2,55] = 4.16, p = .022; please see Figure D2). There was no main effect of expertise (p = .226). To further examine the interaction, Bonferroni-corrected t-tests were performed on clinicians’ and clinical trainees’ judgments separately; these revealed no significant effects for clinicians (all ps ≥ .093). Thus, clinicians did not judge the treatment need of PTSD behaviors differently across event contexts (Question A3). One possibility is that clinicians may have judged PTSD treatment need on the merits of the behavioral symptoms alone, which were the exact same across contexts. Clinical trainees, in contrast, judged PTSD behaviors in A1 traumatic event context (M = 7.86, SD = .90; t[15] = -3.31, p = .005) and in non-A1 event context (marginally, M = 6.75, SD = 1.83; t[16] = -2.14, p =
.048) to be more in need of PTSD treatment than PTSD behaviors in everyday event context ($M = 4.70, SD = 2.63$). Perhaps clinical trainees found evidence of PTSD behaviors combined with a traumatic event, either fitting the *DSM-IV-TR* (APA, 2000) and *DSM-5* (APA, 2013) A1 criteria or not, to describe a more PTSD-like and disorder-like set of behaviors than if an everyday event yielded the same behaviors. Interestingly, clinical trainees found PTSD behaviors to be both more likely to have PTSD and more in need of treatment for PTSD when those PTSD behaviors were presented in traumatic event context.

**Primary analyses: Distress judgments for PTSD behaviors.** Somewhat surprisingly, the analysis of variance for distress judgments for PTSD behaviors revealed no significant main effects, nor an interaction (all $p$s $> .329$). Thus, participants were not influenced by event context when judging the distress level of hypothetical people with PTSD behaviors. Again, participants were asked to consider the behavioral symptoms alone when making these judgments when assessing the distress level of people with PTSD behaviors, and the behavioral symptoms were identical across contexts.

**Primary analyses: Functional impairment judgments for PTSD behaviors.** The ANOVA on functional impairment judgments for PTSD behaviors revealed a main effect of event context ($F[2,50] = 3.25, p = .047$), but Bonferroni-corrected t-tests revealed no significant effects (all $p$s $> .055$), making this effect hard to interpret. The means for functional impairment judgments in A1 traumatic event context ($M = 7.33, SD = .84$) and non-A1 traumatic event context ($M = 7.29, SD = .85$) tend toward running (non-significantly) higher than judgments in everyday event context ($M = 6.48, SD = 1.66$). There was also no main effect of expertise, nor an interaction (all $p$s $> .494$), such that participants were unaffected by context when making a second judgment for PTSD behaviors. Perhaps, people consider any type of traumatic event to
cause functional impairment in the person’s life (e.g., because they were unemployed), even if slight, compared to having merely experienced an everyday event (e.g., listening to a neighbor talk about war).

**Primary analyses: Psychological abnormality judgments for PTSD behaviors.** The ANOVA on psychological abnormality judgments for PTSD behaviors revealed no significant effects (all $p$s $\geq .159$). Thus, we found no evidence that participants were influenced by event context when judging the psychological abnormality of hypothetical people with PTSD behaviors. This is the third PTSD behavior judgment that was not influenced by context. This does not align with our Experiment 1 finding that PTSD behaviors were judged to be less psychologically abnormal in traumatic event context than in everyday event context. However, the means do go in the same direction in both experiments, with PTSD behaviors appearing to be rated as the most psychologically abnormal in everyday event context (in Appendix E, $M = 7.10$, $SD = 1.41$), followed by A1 traumatic event context ($M = 6.33$, $SD = 1.82$) and non-A1 traumatic event context ($M = 6.06$, $SD = 1.75$). A smaller number of participants, resulting in diminished power, and fewer total vignettes rated in this study, contributing to increased variability in ratings, are likely to have contributed to this failure to replicate (272 vignettes were judged in Experiment 1 and 56 in Appendix E).

**Primary analyses: Difficulty understanding judgments for PTSD behaviors.** The ANOVA on difficulty understanding judgments for PTSD revealed a significant effect of event context ($F[2,50] = 9.64$, $p < .001$). Bonferroni-corrected t-tests revealed that participants found PTSD behaviors in both non-A1 traumatic event context ($M = 3.00$, $SD = 1.54$; $t[36] = 3.42$, $p = .002$) and A1 traumatic event context ($M = 3.00$, $SD = 1.41$; $t[37] = 3.59$, $p = .001$) to be less difficult to understand than PTSD behaviors in everyday event context ($M = 5.05$, $SD = 2.04$).
Thus, the context of a traumatic event, fitting either A1 or non-A1 criteria, led people to judge PTSD behaviors to be less difficult to understand compared to everyday event context. There was no main effect of expertise, nor an interaction (all $p > .232$). This replicates our Experiment 1 finding that PTSD behaviors were judged to be less difficult to understand in traumatic event context than in everyday event context, and extends this past work to include non-A1 traumatic event context.

In response to Question A3, whether event context affected non-disorder judgments for PTSD behaviors, we surprisingly found that context only influenced two of five judgments (treatment need and difficulty understanding judgments). To our participants, PTSD seemed more in need of treatment in traumatic event context than in everyday event context. Although this was found for non-A1 traumatic event context in particular, the means for A1 traumatic event compared to everyday event context also go in the same direction. Thus we speculate that even though clinicians may not feel that PTSD behaviors that occur in traumatic event context are particularly psychologically abnormal, they still consider them to be much in need of PTSD treatment, which contributed to their making more PTSD diagnoses in traumatic event context than in everyday event context.

Difficulty understanding is not a consideration necessarily thought to align with diagnosis. Thus, out of all factors we thought might align with PTSD diagnoses (treatment need, distress, functional impairment), only treatment need judgments were similarly influenced by context, and even so, only for clinical trainees. As discussed above, one possibility is that a smaller number of subjects in this study masked the effects of interest. Another possibility is that all three events could have been viewed as sufficiently traumatic; however, these events were normed to differ in perceived degree of trauma (please see p. 216). Possibly, clinicians and
clinical trainees generally do not feel that someone with PTSD behaviors is any more or less distressed, functionally impaired, or psychologically abnormal just because their behavioral symptoms manifest after one type of event compared to another.

Limitations

This study was limited by a smaller number of subjects compared to Experiment 1, and fewer vignettes included for ratings, making comparison between the two studies somewhat difficult. In Experiment 1, 136 people participated and each judged six vignettes, for a total of 816 vignettes judged in Experiment 1. On the other hand, in Appendix E, 56 people participated and each judged three vignettes, for total of 168 vignettes judged in Appendix E. That said, many key findings were replicated across studies, including that PTSD and MDD behaviors were judged to be less difficult to understand in traumatic event context than in everyday event context.

Implications

Much philosophical discussion among clinicians has pointed to considerations of treatment need, distress, functional impairment, and psychological abnormality as ideas tightly linked to the presence of a mental disorder, and accordingly, to the presence of behaviors that merit a mental disorder diagnosis (e.g., Bennett, 2011). Even though clinicians and clinical trainees in Experiment 1 judged PTSD behaviors to be less psychologically abnormal in the context of a traumatic event than in the context of an everyday event, they nevertheless made more PTSD diagnoses for PTSD behaviors that occurred in traumatic event context than in everyday event context. In this experiment, we explored whether judgments of treatment need, distress, and functional impairment – considerations philosophically associated with diagnosis need – better align with PTSD diagnoses than judgments of psychological abnormality. The
results of the current exploratory work suggest that PTSD diagnosis likelihood judgments are predicted by considerations of PTSD treatment need judgments when treatment need, distress, functional impairment, and psychological abnormality were simultaneously considered. Thus, for PTSD behaviors specifically, treatment need and diagnosis need are ideas that are tightly linked together, as evidenced by clinicians’ judgments. This was also the case for MDD behaviors, suggesting a disorder-general effect. However, we don’t know from these data whether judgments of treatment need lead people to view PTSD and MDD behaviors as needing a diagnosis, or whether diagnosis judgments lead people to view PTSD and MDD behaviors as needing treatment. Future systematic experimental work remains to be conducted to investigate this question of causality more directly.

We also asked whether A1 traumatic event context increases PTSD diagnoses relative to everyday event context, to the same extent as do non-A1 traumatic events relative to everyday event context. We found that PTSD diagnosis likelihood judgments were increased by A1 traumatic event context relative to everyday event context, but they were not increased by non-A1 traumatic event context relative to everyday event context. Thus, people’s judgments continued to align with DSM criteria wherein an A1 traumatic event is required for a PTSD diagnosis. If clinicians and clinical trainees continue to diagnose PTSD only in the context required for a PTSD diagnosis in the DSM-5 (APA, 2013), they will adhere to the manual, but they may not end up diagnosing all people who are exhibiting true PTSD behavioral symptoms and might benefit greatly from treatment (that is, true PTSD behavioral symptoms only in a non-DSM-5 [APA, 2013] view of potential PTSD, based on research such as that conducted by Mol et al., 2005). On the other hand, treatment need judgments were increased the most by non-A1 traumatic context. One possibility is that PTSD behaviors that occur in the context of a non-A1
traumatic event are viewed as meriting treatment the most because clinicians recognize this type of event to be traumatic but not currently recognized by the DSM as adhering to a known disorder, perhaps making the case seem more atypical. PTSD is a unique DSM mental disorder because it includes context in its diagnostic criteria. As we are finding, PTSD may also be unique in that clinicians and clinical trainees do not use life event context to explain away PTSD diagnoses, but instead use it to pinpoint behaviors that need treatment for PTSD and therefore merit a PTSD diagnosis.
### Appendix E Dependent Variable Text and Instructions

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Text</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTSD Diagnosis</td>
<td>Given [X]’s subsequent behaviors, how unlikely or likely is it that [X] has Posttraumatic Stress Disorder?</td>
<td>Please choose your response on a scale of 1-9, where 1=very unlikely, and 9=very likely.</td>
</tr>
<tr>
<td>MDD Diagnosis</td>
<td>Given [X]’s subsequent behaviors, how unlikely or likely is it that [X] has Major Depressive Disorder?</td>
<td>Please choose your response on a scale of 1-9, where 1=very unlikely, and 9=very likely.</td>
</tr>
<tr>
<td>Any Diagnosis</td>
<td>Given [X]’s subsequent behaviors, how unlikely or likely is it that [X] has ANY diagnosable DSM disorder?</td>
<td>Please choose your response on a scale of 1-9, where 1=very unlikely, and 9=very likely.</td>
</tr>
<tr>
<td>PTSD Treatment</td>
<td>Given [X]’s subsequent behaviors, is [X] unlikely or likely to need any type of professional mental health treatment for [Posttraumatic Stress Disorder] (e.g., psychotherapy, medication, etc.)?</td>
<td>Please choose your response on a scale of 1-9, where 1=very unlikely and 9=very likely.</td>
</tr>
<tr>
<td>MDD Treatment</td>
<td>Given [X]’s subsequent behaviors, is [X] unlikely or likely to need any type of professional mental health treatment for [Posttraumatic Stress Disorder] (e.g., psychotherapy, medication, etc.)?</td>
<td>Please choose your response on a scale of 1-9, where 1=very unlikely and 9=very likely.</td>
</tr>
<tr>
<td>Category</td>
<td>Description</td>
<td>Rating Scale</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Distress</td>
<td>Given [X]'s subsequent behaviors, how much distress is [X] experiencing?</td>
<td>Please choose your response on a scale of 1-9, where 1 = no distress at all, and 9 = the most distress possible.</td>
</tr>
<tr>
<td>Functional Impairment</td>
<td>Given [X]'s subsequent behaviors, how much functional impairment is [X] experiencing?</td>
<td>Please choose your response on a scale of 1-9, where 1 = no functional impairment at all, and 9 = the most functional impairment possible.</td>
</tr>
<tr>
<td>Psychological Abnormality</td>
<td>How psychologically normal or psychologically abnormal are [X]'s subsequent behaviors?</td>
<td>Please choose your response on a scale of 1-9, where 1 = very psychologically normal, and 9 = very psychologically abnormal.</td>
</tr>
<tr>
<td>Difficulty Understanding</td>
<td>How easy or difficult is it to understand [X]'s subsequent behaviors?</td>
<td>Please choose your response on a scale of 1-9, where 1 = very easy to understand, and 9 = very difficult to understand.</td>
</tr>
</tbody>
</table>
Table D2

*Non-DSM Event Vignette Type and Text*

<table>
<thead>
<tr>
<th>Vignette Name</th>
<th>Type of Non-DSM Event</th>
<th>Vignette Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diana</td>
<td>Burglary</td>
<td>Diana lives in a large suburban town with her husband and two kids, and teaches at the elementary school. One day she arrived home alone. She realized that the TV and computers were gone. The checkbook, cash, and passports were missing from the safe. Diana’s jewelry, including treasured family heirlooms, had been taken. Drawers full of the children’s clothes had been dumped out and trampled with large, dirty footprints. She called the police. Strangely, none of the doors or windows, or even the safe, had been damaged. The police concluded that someone had likely watched the family very closely, stolen a key, and waited for the right moment. It could have been a stranger or someone close to the family. (101 words)</td>
</tr>
<tr>
<td>Martin</td>
<td>Divorce</td>
<td>Martin lives with his wife and two children whom he often drives to school. He arrived home one day to find his wife packing all of her belongings into a truck. She announced that she was leaving Martin, and presented him with divorce papers. He tried to talk her out of the divorce, but only caused painful arguments that he knows his children overheard. She would not reveal where she went. Martin learned through his lawyer that his wife had also</td>
</tr>
</tbody>
</table>
filed for sole custody of the children, so he might only see his kids on holidays. Martin soon discovered that his wife was already living with someone new, in a town that was only two hours away.

(104 words)

Justin | Unemployment | Justin worked at a shipping company in a large suburban town, supporting his wife and two kids. On one occasion, Justin’s co-worker made an expensive ordering mistake and falsely blamed it on Justin. Justin was fired that day. His boss did not give an explanation, and forced Justin to leave immediately. After, he was not able to find other work because his boss refused to serve as a reference. When Justin ran out of money, he was forced to move his family to a homeless shelter. (69 words)

*Note.* Following the format of Appendix A, underlining indicates background information. Word count corresponds to the event description, excluding background information.
Table D3

*PTSD Behaviors Correlation Matrix for PTSD Diagnosis Likelihood Judgments, PTSD Treatment Need Judgments, Distress Judgments, Functional Impairment Judgments, and Psychological Abnormality Judgments*

<table>
<thead>
<tr>
<th></th>
<th>PTSD diagnosis likelihood</th>
<th>PTSD treatment need</th>
<th>Distress</th>
<th>Functional impairment</th>
<th>Psychological Abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTSD</td>
<td><strong>.662</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distress</td>
<td>.254</td>
<td>.303</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional</td>
<td>.260</td>
<td>.397</td>
<td>.530</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological</td>
<td>-.048</td>
<td>.140</td>
<td>.182</td>
<td>.113</td>
<td></td>
</tr>
</tbody>
</table>

**p < .001  
*p ≤ .05
Table D4

*Multiple Regression Results for PTSD Behaviors*

<table>
<thead>
<tr>
<th>Predictors</th>
<th>PTSD treatment need</th>
<th>Distress</th>
<th>Functional impairment</th>
<th>Psychological abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>β</strong></td>
<td>.609 (5.88)**</td>
<td>.291 (.815)</td>
<td>-.068 (-.324)</td>
<td>-.200 (-1.47)</td>
</tr>
</tbody>
</table>

Note: t values shown in parentheses, **p < .001.**
Table D5

*MDD Behaviors Correlation Matrix for MDD Diagnosis Likelihood Judgments, MDD Treatment Need Judgments, Distress Judgments, Functional Impairment Judgments, and Psychological Abnormality Judgments*

<table>
<thead>
<tr>
<th></th>
<th>MDD diagnosis</th>
<th>MDD treatment</th>
<th>Distress</th>
<th>Functional impairment</th>
<th>Psychological abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDD diagnosis</td>
<td>.642**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDD treatment</td>
<td>.548</td>
<td>.581**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distress</td>
<td></td>
<td></td>
<td>.473**</td>
<td>.430**</td>
<td>.669**</td>
</tr>
<tr>
<td>Functional impairment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological abnormality</td>
<td>.325</td>
<td>.296</td>
<td>.288*</td>
<td>.316*</td>
<td></td>
</tr>
</tbody>
</table>

** $p < .001$

* $p \leq .05$
Table D6

*Multiple Regression Results for MDD Behaviors*

<table>
<thead>
<tr>
<th>Predictors</th>
<th>MDD treatment need</th>
<th>Distress</th>
<th>Functional impairment</th>
<th>Psychological abnormality</th>
</tr>
</thead>
<tbody>
<tr>
<td>β</td>
<td>.451 (3.62)**</td>
<td>.187 (1.05)</td>
<td>.168 (.984)</td>
<td>.059 (.907)</td>
</tr>
</tbody>
</table>

Note: *t* values shown in parentheses, ** *p* \( \leq .001.\)
Figure D1

PTSD Diagnosis Likelihood Judgments for PTSD Behaviors by Event Type And Expertise
Figure D2

PTSD Treatment Need Judgments for PTSD Behaviors by Event Type And Expertise