Proportionate Responses to Life Events Influence Clinicians' Judgments Of Psychological Abnormality

Nancy S. Kim,
Department of Psychology, Northeastern University

Daniel J. Paulus,
Department of Psychology, Northeastern University

Jeffrey S. Gonzalez, and
Ferkauf Graduate School of Psychology, Yeshiva University, and Diabetes Research Center, Albert Einstein College of Medicine

Danielle Khalife
Department of Psychology, Northeastern University

Abstract

Psychological abnormality is a fundamental concept in the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV-TR*; APA, 2000) and in all clinical evaluations. How do practicing clinical psychologists use the context of life events to judge the abnormality of a person's current behaviors? The appropriate role of life-event context in assessment has long been the subject of intense debate and scrutiny among clinical theorists, yet relatively little is known about clinicians' own judgments in practice. We propose a *proportionate-response* hypothesis, such that judgments of abnormality are influenced by whether the behaviors are a disproportionate response to past events, rendering them difficult to understand or explain. We presented licensed, practicing clinical psychologists (*N*=77) with vignettes describing hypothetical people's behaviors (disordered, mildly distressed, or unaffected) that had been preceded by either traumatic or mildly distressing events. Experts' judgments of abnormality were strongly and systematically influenced by the degree of mismatch between the past event and current behaviors in strength and valence, such that the greater the mismatch, the more abnormal the person seemed. A separate, additional group of clinical psychologists (*N*=20) further confirmed that the greater the degree of mismatch, the greater the perceived difficulty in understanding the patient. These findings held true across clinicians of different theoretical orientations and in disorders for which these patterns of judgments ran contrary to formal recommendations in the *DSM-IV-TR* (APA, 2000). The rationality of these effects and implications for clinical decision science are discussed.

Keywords

assessment; clinical judgments; explanation; causal reasoning

---

Correspondence concerning this article should be directed to Nancy S. Kim, Department of Psychology, 125 Nightingale Hall, Northeastern University, 360 Huntington Avenue, Boston, Massachusetts 02115-5000, USA. n.kim@neu.edu.
Daniel J. Paulus is now at Department of Psychology, University of Massachusetts Boston.

Proportionate Responses to Life Events Affect Clinicians' Judgments Of Psychological Abnormality

What factors influence how clinicians judge the abnormality of people's behaviors? Among the myriad factors converging to influence such judgments are not only data from formal measures, such as clinical observation of the client's presenting symptoms, but also socio-demographic descriptors of the client such as gender, race, and class (e.g., Kessler & Neighbors, 1986; Lincoln, 2006; Rosenfield, 1982), and variables known to influence cognitive processing such as the availability of diagnoses given to previously seen patients, the terminology used to describe the patients, the range and variability of prior patients with the same diagnosis stored in memory (e.g., Hatala, Norman, & Brooks, 1999; Norman, Arfai, Gupta, Brooks, & Eva, 2003; Young, Brooks, & Norman, 2007), and the clinician's own causal and explanatory reasoning (e.g., Kim & Ahn, 2002; Ahn, Proctor, & Flanagan, 2009; de Kwaadsteniet, Hagmayer, Krol, & Witteman, 2010; Meehl, 1973).

In the current paper, we focus on the last of these factors. Specifically, we ask how causal and explanatory reasoning about current symptoms may systematically influence clinicians' judgments of psychological abnormality. The concept of *psychological abnormality* has long been, and continues to be, a central concept in the *DSM* (APA, 2000) and in a wide range of clinical evaluations. This is reflected, for example, in how the term is used to characterize the field. The introduction to the *DSM-IV-TR* (APA, 2000) itself explicitly states that the provided operational definition of mental disorder is intended to help guide "decisions regarding which conditions on the boundary between normality and pathology should be included in *DSM-IV* (p. xxxi);" furthermore, "Abnormal Psychology" is the title of nearly every student of psychology's initial introduction to the concept of mental disorders. In other work, we have shown that judgments of psychological abnormality strongly predict diagnostic judgments, estimates of statistical rarity, judgments of cultural unacceptability, and judgments of maladaptiveness in daily functioning (Kim, Nguyen, & Paulus, 2011; Yopchick & Kim, 2011). For these reasons, it is important to gain a better understanding of what practicing clinicians believe constitutes psychological abnormality.

Judgments of psychological abnormality most directly speak to the fundamental issue of uncovering clinicians' true beliefs about what constitutes psychopathology. That is, judgments of psychological abnormality should most straightforwardly measure what clinicians believe to be abnormal or psychopathological, whereas a *DSM* diagnostic judgment likely measures a combination of knowledge of the manual, the ability and willingness to apply recommended diagnostic criteria, the additional consideration of matching the case to a diagnostic category, and so on. Indeed, in his seminal work on reflective practice, Schön (1983) highlights how practitioners can easily recognize deviations from the norm, even when they are unable to verbalize what is normal, instead identifying "bad fits" to familiar patterns. As our work above suggests, clinicians' basic beliefs about abnormality may have a permeating influence across the range of judgments that they make in clinical practice. For example, clinicians do make decisions about which *DSM* diagnosis to give, but they won't (or certainly shouldn't) assign a *DSM* diagnosis if they believe the client is *not* in fact exhibiting psychopathology. That is, abnormality judgments dictate the degree to which a specific diagnostic judgment is deemed necessary to make in the first place. Additionally, research has shown that clinicians often use "workarounds" in practice, essentially circumventing the authority of the *DSM* to allow themselves room to apply their own clinical judgment (Whooley, 2010). For this reason, we have chosen to examine judgments of psychopathology to more directly observe how clinicians truly judge psychopathology without the aforementioned workarounds. In sum, we focus not on formally structured manual-based judgments, but rather on clinicians' more

fundamental judgments concerning what constitutes psychopathology. Thus, clinicians' judgments of abnormality are the focus of our study.

Our central aim is to propose and test a *proportionate-response hypothesis*, such that clinicians' judgments of abnormality are influenced by whether a person's behaviors are perceived to be a disproportionate response to past events, rendering them difficult to understand or explain. Specifically, we hypothesize that if the disproportionality of the person's response renders it difficult to understand how the life event could have caused the behaviors, then the person will be judged more abnormal. If the response is proportionate to the event, making it easy to understand that causal link, then the person will be judged less abnormal. Wakefield (1997) suggested that this rule of proportionality helps to define the limits of what constitutes "normal" reactions to life stressors. By understanding what the appropriate or culturally sanctioned responses are, we can see that, for example, if distressed responses are of increased intensity or extended duration, they may be inappropriate or disproportionate to their cause, indicating dysfunction or psychopathology.

We tested this hypothesis by manipulating the degree of mismatch in negativity between the life event and behaviors in hypothetical cases, separately testing to ensure that a mismatch made it more difficult to understand the subsequent behaviors of that person than a match. Seminal work in cognitive science has indicated that in general, it is cognitively difficult to draw a causal link between a strong cause and weak effect, or between a weak cause and strong effect (Einhorn & Hogarth, 1986; Medin, 1989). In contrast, when the two are matched in strength, it is much easier to mentally construct a causal link between them. In past work showing an influence of causal information on abnormality judgments (Ahn, Novick, & Kim, 2003; Kim & LoSavio, 2009; Meehl, 1973), the causal mechanisms were described in detail for participants and it was therefore always easy for them to draw the causal link; in addition, the cause and effect were always relatively well matched in strength.

For example, in one prior study, undergraduates and expert and novice clinical psychologists read vignettes describing hypothetical clients, such as a woman who frequently has insomnia, has trouble remembering the names of objects, and has episodes of extreme anxiety. Across expertise levels, people who did not receive an explanation for why the woman had these symptoms in the first place rated her as quite abnormal. In contrast, those who were given a plausible causal mechanism for what precipitated those symptoms (e.g., she is very stressed due to her workload, which causes her to frequently suffer from insomnia and be in a habitual state of sleep deprivation) rated her as significantly more normal than people who were not given the explanation (Ahn et al., 2003). More recently, a parallel effect was also found, again with hypothetical client vignettes in which the cause and effect are relatively well matched in strength, in undergraduates who were asked to make judgments of people's need for psychological treatment when given an explanation versus not (Kim & LoSavio, 2009).

However, in real life, it is unlikely that the details of the causal mechanism linking past events to current symptoms will be described explicitly for clinicians; instead, it is more likely that they will glean various pieces of information about the client that they would link together only by inference (Pennington & Hastie, 1988). To test our currently proposed *proportionate-response hypothesis*, it was essential to examine cases in which the strength of the cause and effect are mismatched, and to allow respondents to attempt to generate the causal mechanism linking the two. The *proportionate-response hypothesis* would predict that mismatched cases should be judged more abnormal, because clinicians will experience more cognitive difficulty in trying to understand the link between the cause and its disproportionate effect. This idea is derived from Kahneman & Tversky's (1982) simulation heuristic, whereby people believe the more easily they can explain an event, the more likely

it is to occur; conversely, struggling to explain an event may also make it seem more deviant from the norm. Again, we suggest that one type of scenario likely to elicit such a cognitive struggle is one in which the cause of the behaviors does not reasonably match the behaviors themselves in strength and in valence.

For example, suppose that a man named "Justin" is no longer interested in anything he used to enjoy, and has trouble falling asleep or staying asleep without medication. A few years later, Justin is feeling intense, daily guilt and generally believes he is worthless, and his insomnia has not improved. According to the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV-TR*; APA, 2000), Justin is showing signs of major depressive disorder (MDD), and therefore is more abnormal than a person who is not experiencing these symptoms and who appears content. However, consider whether our assessment of Justin's abnormality is affected by knowing exactly what kind of event precipitated these symptoms. For instance, suppose we learn that a traumatic event caused his MDD symptoms (e.g., several years ago, he experienced active combat duty, during which he killed a civilian). Or, suppose that Justin experienced no traumatic event firsthand, but rather learned that a similar event had happened to an acquaintance, and then experienced the exact same intense depressive symptoms. In the first case, the symptoms seem comparatively better justified by his life experiences and the causal story is more clear; in the second case, it is more difficult to explain why he would experience depressive symptoms following such a relatively benign precipitant. Thus, the proportionate-response account would predict that in the first case Justin would seem relatively less abnormal than in the second case, despite having the same symptoms. In the same vein, if Justin shows very contented, unperturbed behaviors immediately following the traumatic event above, and continues to do so for a period of time, he may be perceived as abnormal relative to the case in which a benign precipitant leads to the same behaviors. This would be the prediction of the proportionate-response account, despite the fact that he is not showing any *DSM*-disordered symptoms in either of those cases.

An alternative hypothesis is a *current-symptom hypothesis*, such that abnormality judgments depend instead on the person's current state; for example, his or her symptoms, distress, and impairment in daily life. That is, people may believe that Justin's current state is what defines him as relatively abnormal versus normal. If he has debilitating, depressive symptoms, people will believe he is abnormal; if he has no symptoms, then they will believe he is normal. Knowing additional information about the nature of the cause will not mitigate abnormality of depressive symptoms, nor will it suggest any underlying dysfunction in asymptomatic cases. This account predicts that people are primarily influenced by current psychological distress and impairment per se in assessing abnormality. The current *DSM-IV-TR* (APA, 2000) explicitly adheres to this approach; for the vast majority of disorders, it was deliberately structured to direct attention to observable symptoms and to attempt to be agnostic or neutral regarding the etiology of disorders (a notable exception is post-traumatic stress disorder [PTSD], discussed in detail below).

Accounting for prior instruction and training

Our aim was to measure clinicians' straightforward judgments of the abnormality of behaviors. Because we wanted to use actual disorders as specified in the *DSM-IV-TR* (APA, 2000), as opposed to the artificial disorders used in past work, it was important for us to account for effects of prior instruction and training. We did so in two different ways: (1) by carefully targeting our disorder selection and (2) by tracking clinicians' theoretical orientations.

1. *DSM-supported versus unsupported causal context.* First, it was necessary to control for explicit instructions given in the *DSM* for how to consider causal

context in assessment. The *DSM*'s overarching philosophy is to operate without reference to any specific theories of causal etiology (APA, 2000). Nonetheless, we identified and focused on two disorders for which the current *DSM* does explicitly give more specific instructions on how to treat precipitating life events in assessment. The *DSM-IV-TR* (APA, 2000) specifies that a traumatic life event is requisite for a diagnosis of PTSD. Therefore, if clinicians make judgments in accord with *DSM* specifications, then any hypothetical clients with PTSD symptoms should still be perceived as abnormal even if a traumatic life event occurred (that is, given that trauma is one of the diagnostic requirements for PTSD in the *DSM*, the presence of a negative life event should not make it less abnormal according to the *DSM*). Conversely, the *DSM* specifies that in diagnosing MDD, most traumatic life events should *not* be factored into judgments.1 Thus, if clinicians follow the *DSM*, then perceptions of the hypothetical clients with MDD symptoms should not change regardless of whether a precipitating life event was traumatic in nature. Alternatively, if we instead find that precipitating life events influence judgments of abnormality in both disorders in accord with the *proportionate-response hypothesis*, this would imply that the effect of precipitating events is strong enough to override prescribed *DSM* specifications.

2. *Psychoanalytic versus cognitive-behavioral theoretical orientations*. Similarly, we also needed to monitor whether the results were the same across clinicians with different theoretical orientations. We reasoned that if theoretical orientation had any moderating effect at all on the key hypothesized outcome, it might appear most clearly in the comparison between psychoanalytic clinicians and cognitive-behavioral clinicians. For example, psychoanalytic clinicians, who traditionally place great importance on the past and how it leads to the present (Shedler, 2010), might be relatively more strongly influenced by precipitating events and their relation to current behaviors than cognitive-behavioral clinicians, who may not be as primarily focused on the past (e.g., Beck, 1995; Clark & Beck, 1999). If, on the other hand, the influence of precipitating events on judgments of abnormality is strong enough to override adherence to explicitly held theoretical approaches, then the hypothesized effect of precipitating life events should not vary across theoretical orientations.

Method

Participants

Main study—Seventy-seven licensed clinical psychologists participated in the main study and were emailed a $30 gift card from a major online retailer as a token of thanks for their time. We mailed postcard invitations to participate to 400 clinical psychologists, all of whom had completed Ph.D. degrees, were currently licensed, and were currently practicing in the U.S. (19.3% response rate). To ensure that the sample was nationally representative, we also attempted to sample clinicians proportionately to the relative population sizes of their respective states. Within these parameters, the 400 clinicians were sampled as follows: 270 clinical psychologists were taken from the American Psychological Association's national directory and contacted with formal permission from the association, an additional 30 were randomly selected from the American Psychological Association's *Directory of Ethnic Minority Professionals in Psychology* (2001), and the remaining 100 were randomly

---

[1]The *DSM-IV-TR* (APA, 2000) excludes cases from a diagnosis of a major depressive episode for bereavement (up to 2 months) and effects of a substance or general medical condition. We therefore carefully constructed our stimulus materials so that there are no *DSM-IV-TR* exclusions for the event described in our example of "Justin" or for any of the precipitating events for depressive symptoms used in our study.

sampled from the American Psychoanalytic Association directory to ensure adequate sampling of psychoanalytically oriented clinicians. We verified the current licensure status of clinicians via state public health department websites prior to recruitment.

The 77 clinical psychologists who opted to participate in the main study had a mean of 17.2 years of experience ($SD = 8.9$) and reported spending a mean of 18.9 hours per week seeing patients ($SD = 10.6$) over their careers. Forty identified their primary theoretical orientation as cognitive, behavioral or cognitive-behavioral (referred to collectively as *cognitive-behavioral*, henceforth), 16 psychoanalytic or psychodynamic (referred to collectively as *psychoanalytic*, henceforth), 9 eclectic, 4 humanistic, 4 family systems, and 4 "other." Clinicians estimated the mean age of their patients to be 30.2 years old ($SD = 12.2$), and reported that 84.6% of their patients had Axis I diagnoses ($SD = 23.7\%$) and 27.3% Axis II diagnoses ($SD = 24.1\%$). There were 42 female and 34 male clinicians (one chose not to disclose gender). Clinicians also reported ethnicity and race. Overall, 15.6% of these participants identified themselves as belonging to a racial or ethnic minority group. With respect to race, 63 self-identified as White, 5 Asian, 2 African-American, and 1 reported multiple races (6 chose not to disclose race). Four self-identified their ethnicity as Hispanic or Latino.

Manipulation check—In a manipulation check of our materials, we also ran a separate, additional group of 20 clinical psychologists meeting the same criteria as above (licensed, currently practicing clinical psychologists with completed Ph.D. degrees). We mailed postcard invitations to participate to 100 clinical psychologists (20.0% response rate) whose current licensure status we verified via state public health department websites prior to recruitment. Clinicians participating in this manipulation check study were each emailed a $25 gift card from a major online retailer as a token of thanks.

The 20 clinical psychologists who participated in the manipulation check had a mean of 26.6 years of experience ($SD = 9.0$) and had spent a mean of 21.9 hours per week seeing patients ($SD = 10.8$) over their careers. Nine self-identified as cognitive-behavioral, 3 as psychoanalytic, 2 as humanistic, 3 as eclectic, and 2 as "other." The mean age of their patients was 36.8 years ($SD = 11.1$), and 87.1% of their patients had Axis I diagnoses ($SD = 14.3\%$) and 14.7% Axis II diagnoses ($SD = 9.5\%$). Seven were female and 12 male (one did not disclose gender). Race was self-reported by 19 participants as White; one did not disclose race. Two participants self-identified as Hispanic/Latino.

Materials and procedure

Each participant was presented with six fabricated vignettes written to mimic the structure, narrative style, and severity of MDD and PTSD cases in the *DSM-IV-TR Casebook* (Spitzer, Gibbon, Skodol, Williams, & First, 2001). Each vignette was comprised of two full paragraphs. In each vignette, the first paragraph described a precipitating event for that client.2 Two precipitating events were written for each vignette; one was traumatic in nature, and the other was a mildly distressing control version of the traumatic event. Traumatic and control events were approximately matched for length. Control versions were created either by changing the story so that the main character was reading or hearing about the traumatic event instead of experiencing it (three vignettes) or by modifying the events themselves so that they were similar, but much more mundane and mild stressors (three vignettes).3 An example of the first type of modification is the case of "Justin," in which the

---

2In this paper, we refer to hypothetical people being assessed for abnormality as "clients" to clearly distinguish them from the study participants. However, in our study materials, we always referred to these hypothetical people using proper names (e.g., "Justin," "Cathy"), and we never called them *clients* or *patients*.

traumatic event was described as having killed a civilian during his tour of duty in Iraq, and in which the control event was described as hearing an acquaintance describe the same event, which had happened to the acquaintance, not to Justin. To give an example of the second type of modification, in the case of "Diana," the traumatic event version described a scenario in which her best friend was in a coma after a severe car accident (among other events), whereas the control event version described a scenario in which her best friend was out of town for the week and therefore unable to interact with Diana as she usually did (among other events similarly downgraded in intensity and negativity). Importantly, event type was manipulated *between* participants, such that they were randomly assigned to either view all six vignettes with traumatic events or all six vignettes with control events.

Each vignette also contained a second paragraph entitled "[vignette character name]'s subsequent behaviors" describing the person's behaviors following the precipitating event. We created three types of behavior descriptions for each vignette: unaffected, mildly distressed, and disordered (i.e., *DSM*-defined symptom criteria of MDD or PTSD), also matched for length ($M = 65$ words). In the unaffected version, the person was described as continuing life exactly as s/he had before the event, feeling unaffected immediately afterwards and up through the current day. The mildly distressed version described the individual as mildly affected by the event, having, for example, a few bad dreams or negative feelings that subsided with the help of support structures, such that s/he was feeling relatively contented and positive at the current time. This mildly distressed version was designed to act as the control case, following a seminal line of research showing that cases describing low-grade distress are the most appropriate control for disordered cases (Schwartz & Link, 1989). In the disordered version, the negative effects were described as greatly distressing, persistent, and clearly interfering with current daily life immediately afterwards and up through the current day.

Three of six vignettes were based on PTSD cases; the other three were based on MDD cases. Behavior type was manipulated within-participant and across vignettes, such that each participant viewed two vignettes with disordered behaviors (one MDD and one PTSD), two with mildly distressed behaviors (one modified from its corresponding MDD case and one from its corresponding PTSD case), and two with unaffected behaviors. Importantly, behavior type was rotated across vignettes between participants, such that each vignette took equal turns being depicted with *each* of the six possible combinations of event type (traumatic, control) and behavior type (unaffected, mildly distressed, disordered).

Participants completed the study online. The vignettes were presented sequentially, in randomized order, to each participant. The key dependent measure for each vignette was a perceived abnormality rating ("How normal or abnormal are [vignette character name]'s subsequent behaviors?" where 1= very normal behaviors and 9= very abnormal behaviors). Participants were able to change each rating while still viewing the vignette page, but once they clicked a separate button to formally submit the rating, they were not allowed to change previous ratings. All participants were informed of this feature in the instructions at the beginning of the experiment. Participants also completed a second task unrelated to the purpose of the current paper; the order of these tasks was counterbalanced between participants, and order analyses showed that this second task had no effect on the results of the abnormality judgment task (see Footnote 4).

---

[3]Follow-up analyses showed that, as intended, the results did not differ across these two types of controls, so we will not discuss it further.
[4]There were no effects or interactions involving Task Order (all p's > .257; all η^2 < .02), so this factor will not be mentioned further.

Then, participants completed a demographics questionnaire reporting gender, ethnicity, and race, with the option to decline to answer any question. In addition, participants answered questions adapted from Lehman (1992) about their primary theoretical orientation, licensure and highest degree, their years of experience actively practicing, average hours per week seeing clients, and current caseload (the average age of their clients, percentage of clients with Axis I diagnoses, and percentage of clients with Axis II diagnoses). Participants were shown a summary of the hypotheses, design, and purpose of the study immediately upon completion of the task.

Manipulation check—We also independently ran a manipulation check study to test the assumption that proportionate responses were easier to understand than disproportionate responses. We presented the separate group of manipulation check participants with almost the exact same materials and design as in our main study, except that after each vignette, they were instead asked an ease-of-understanding question ("Given this information, how easy or difficult is it to understand [vignette character's name] current behaviors?" where 1= very easy to understand and 9= very difficult to understand).

Results and Discussion

Ease-of-Understanding Manipulation Check

Again, the *proportionate-response hypothesis* is that behaviors are perceived to be more abnormal when the life-event cause and resulting behaviors are mismatched in strength, forcing clinicians to cognitively struggle to understand the causal link between them. First, we examined our manipulation check results to verify our assumption that cause-effect mismatches decrease the ease with which clinicians are able to understand the behaviors. As expected, a 2 (Event Type; Traumatic, Control) X 3 (Behavior Type; Unaffected, Mildly Distressed, Disordered) ANOVA conducted on the manipulation check data revealed a two-way interaction, $F(2, 36) = 26.9$; $p < .001$; $\eta^2 = .60$. In accord with our assumption, it was much easier for clinicians to understand matched than mismatched cases: Unaffected behaviors were much more understandable following a control event ($M = 2.3$; $SD = 1.5$) than following a traumatic event ($M = 6.5$; $SD = 1.9$; $t[18] = 5.5$; $p < .001$; $\eta^2 = .63$), and disordered behaviors were much more understandable following a traumatic event ($M = 2.4$; $SD = 1.1$) than following a control event ($M = 6.3$; $SD = 1.5$; $t[18] = 6.6$; $p < .001$; $\eta^2 = .71$). Not surprisingly, ease-of-understanding ratings for mildly distressed behaviors did not differ between event types ($p = .375$; $\eta^2 = .04$). See Figure 1b. In sum, the proportionality of responses affected ease-of-understanding exactly as we presumed.

Abnormality Judgments: Main Study

Influences of cause-effect mismatches—Next, we examined the main study data to test whether abnormality judgments were affected by cause-effect mismatches, as predicted by the *proportionate-response hypothesis*. In our study, mismatched cause-effect pairs were the unaffected behaviors following a traumatic event and disordered behaviors following a control event. A 2 (Event Type; Traumatic, Control) X 3 (Behavior Type; Unaffected, Mildly Distressed, Disordered) X 2 (Symptom Type; MDD, PTSD) X 2 (Task Order; Abnormality Task first, Unrelated Task first)4 ANOVA revealed the critical two-way interaction of Event Type and Behavior Type, $F(2, 146) = 65.2$; $p < .001$; $\eta^2 = .47$. See Figure 1a. Unaffected behaviors were perceived to be significantly more abnormal following a traumatic event ($M = 4.4$; $SD = 1.9$) than following a control event ($M = 1.5$; $SD = 0.8$; $t[75] = 8.9$; $p < .001$; $\eta^2 = .51$). Similarly, disordered behaviors following a control event ($M = 7.6$; $SD = 1.1$) were perceived as significantly more abnormal than the same behaviors following a traumatic event ($M = 5.3$; $SD = 1.9$; $t[75] = 6.5$; $p < .001$; $\eta^2 = .36$). Event Type did not affect judgments about the abnormality of mildly distressed behaviors ($p = .340$; η^2

= .01), just as expected given that it did not affect ease-of-understanding of mildly distressed behaviors in the manipulation check study. These findings provide evidence for one part of the *proportionate-response hypothesis*, insofar as they show that behaviors are perceived to be more abnormal when the life-event cause and resulting behaviors are mismatched in strength. The other part of the *proportionate-response hypothesis*, that cause-effect mismatches force people to cognitively struggle to understand the causal link between them, was tested and supported in the manipulation check (see above). There was a main effect of Behavior Type (F[2, 146] = 156.2; $p < .001$; η^2 = .68); however, this was qualified by the above two-way interaction. Finally, there was no main effect of Event Type ($p = .637$, $\eta^2 < .01$).

Influence of DSM specifications—We found that the hypothesized two-way interaction of Event Type and Behavior Type above was qualified by a three-way interaction of Event Type, Behavior Type, and Symptom Type (F[2, 146] = 4.6; $p = .011$; η^2 = .06). Specifically, the influence of cause-effect mismatches on abnormality judgments was dependent upon the disorder under consideration (i.e., MDD versus PTSD). (See Figure 2.) However, post-hoc analyses stratifying by disorder revealed the same two-way interaction of Event Type and Behavior Type for each disorder (for MDD; F[2, 150] = 40.9; $p < .001$; η^2 = .35; for PTSD; F[2, 150] = 48.8; $p < .001$; η^2 = .39). As can be seen in Figure 2, ratings were almost identical for MDD versus PTSD except in the case of disordered behaviors. Although PTSD-disordered behaviors and MDD-disordered behaviors were seen as equally abnormal when paired with control events (PTSD: M = 7.7; SD = 0.8; MDD: M = 7.5; SD = 1.9; t[38] = .7; p = .497; η^2 = .01), people perceived the disordered PTSD vignettes as less abnormal than the disordered MDD vignettes when paired with a traumatic event (PTSD: M = 4.5; SD = 2.4; MDD: M = 6.1; SD = 2.3; t[38] = 3.8; $p < .001$; η^2 = .28). In other words, proportionate responses to trauma decreased perceptions of abnormality more strongly for PTSD than MDD vignettes. This difference appears to have driven the three-way interaction above. Furthermore, all of the key cause-effect mismatch effects were found across the board for *both* disorders, analyzed separately in pairwise comparisons (all p's   .003; all η^2 .11).

To summarize the results in this subsection, abnormality judgments did not, overall, reflect adherence to the *DSM*. First, the PTSD ratings run contrary to that predicted by adherence to the *DSM*, such that disordered behaviors were perceived as less abnormal when precipitated by a traumatic event, despite the fact that the *DSM* specifies that traumatic events should serve as additional evidence of PTSD. In addition, clinicians did not ignore the non-bereavement life events in their judgments about MDD vignettes, as one might expect if clinicians strictly followed *DSM* specifications for MDD. Instead, in both PTSD and MDD vignettes, clinicians' judgments about abnormality were strongly influenced by how well the cause and effect matched in valence and strength (in addition to the presence of the symptoms themselves). Taken together with the results of our manipulation check, these findings speak to the strength of the proportionate-response effect on clinicians' judgments about abnormality.

Influence of clinicians' theoretical orientations—We also conducted a 2 (Event Type) X 3 (Behavior Type) X 2 (Orientation; Psychoanalytic, Cognitive-Behavioral) ANOVA on the abnormality ratings from the clinicians identifying themselves as either psychoanalytic or cognitive-behavioral, and found no effects or interactions involving Orientation (all F 1.0, all p .358, all η^2 .02). Thus, clinicians of both of these theoretical orientations showed effects of cause-effect mismatches on judgments of abnormality. There were too few clinicians in each of the remaining theoretical orientation groups to run analyses on them separately, but we lumped them together into an "Other" theoretical orientation group to ensure that they too did not deviate from the main pattern of

results. A 2 (Event Type) X 3 (Behavior Type) X 3 (Orientation; Psychoanalytic, Cognitive-Behavioral, Other) ANOVA on the abnormality ratings also showed no effects or interactions involving Orientation when this third group was included (all F 1.3, all p . 271, all $\eta^2 < .04$).

Summary—Precipitating life events strongly influenced clinicians' abnormality judgments, regardless of whether the *DSM-IV-TR* specifies that life-event causes for the disorder in question should be taken into account or ignored, and regardless of clinicians' theoretical orientations. In making abnormality judgments, clinicians were influenced by the degree to which the life-event cause matched the current behaviors in strength and in valence; mismatches led to judgments of greater abnormality, even when the hypothetical clients were not experiencing any DSM symptoms. Not surprisingly, clinicians also took into account the nature of the behaviors themselves (i.e., their severity). Overall, these findings indicate that clinicians systematically use causal information in making assessments of the abnormality of human behavior and that this finding cannot be explained away by adherence to prior instruction or training.

General Discussion

Historical accounts of assessment indicate that over a long history of thought, it was the norm to consider causal context in assessing abnormality. Approaches dating back to Aristotle and Kraepelin (Carson, 1996; Horwitz & Wakefield, 2007) and up through the *DSM-II* (APA, 1968) incorporated this notion. With the *DSM-III* (APA, 1980) came an enormous, deliberate restructuring of the diagnostic system to center around making diagnoses by observable symptoms, largely independently of causal context or etiology (Kihlstrom, 2002). Aimed at increasing the consistency of diagnoses within and between clinicians, these modern versions of the *DSM* took "a descriptive approach that attempted to be neutral with respect to theories of etiology" (APA, 1994; pp. xvii–xviii). Given the recency of this shift, a key question is how clinicians would elect to deal with such an etiology-neutral diagnostic system; this speaks to the issue of the *DSM* system's clinical utility or usefulness in practice. Our results suggest that experienced clinicians are likely to sidestep the current *DSM* system's guidance, and instead make judgments consistent with a much older history of understanding abnormality as not only a function of degree of symptom severity, but also with reference to whether there is a reasonable cause for those symptoms.

Specifically, the current study shows that in judging abnormality, clinicians are influenced by the degree of mismatch, in strength and valence, between current behaviors and their life-event precipitants. In our study, a life event that was mismatched to a set of behaviors in strength and valence led to ratings of greater abnormality relative to a life event that was well matched to the same behaviors. This finding held true whether the behaviors themselves were disordered versus unperturbed in nature. Thus, even everyday behaviors can be perceived by clinicians to be quite abnormal, depending on the causal context. Moreover, our manipulation check study showed that clinicians found the mismatched versions of vignettes to be much more difficult to understand than the matched versions. Taking these results together, we propose that abnormality judgments are influenced by whether the behaviors constitute a proportionate response to past events, making it easier to understand how past events led to current behaviors.

Previous research showed that knowing how a life event precipitated disordered behaviors leads people to see those behaviors as more normal, but only tested for this effect using cases for which the life event and behaviors were relatively well matched in strength and valence. Thus, the current project is the first to propose and fully test a *proportionate-*

response hypothesis, such that the inverse also holds true for mismatched cases: When it is *difficult* to explain how traumatic past events could reasonably have led to a set of behaviors because the response is disproportionate to the event, those behaviors are perceived as much *less* normal. Indeed, clinical theorists such as Lifton (1988) have asserted that PTSD symptoms are in fact a "normal adaptive process of reaction to an abnormal situation," which in turn implies that failing to respond to trauma in a PTSD-like way, or showing PTSD symptoms to non-traumatic events, could be construed as abnormal (Gold, Marx, Soler-Baillo, & Sloan, 2005; Seides, 2010; Smith, 2004). Similarly, depression as a response to stressful environmental situations, such as being continuously subjected to sexual harassment at the workplace, has been argued to be a perfectly reasonable response that should be recognized by culturally competent clinicians, whereas not reacting to the same situation should be cause for concern (Zalaquett, Fuerth, Stein, Ivey, & Ivey, 2008).

From a cognitive point of view, this approach makes sense if people represent the client case as a causal system (Ahn et al., 2009; de Kwaadsteniet et al., 2010; Yopchick & Kim, 2009), and if they view the life event within that system as a strong potential cause of future undesirable behaviors. Our work showed that clinicians seem to believe that if a traumatic event has occurred, as a strong cause it would be reasonably expected to give rise to disorder symptoms. In fact, remarkably, clinicians in our study felt that wholly unperturbed behaviors were no less abnormal than full-blown *DSM* disordered behaviors if the person had experienced a traumatic life event (p's   .1; $\eta^2 < .08$). These findings suggest that the influence of life-event context is far from trivial.

Furthermore, two additional sets of analyses and findings suggested that the above basic findings reflect strong beliefs about how life-event context is relevant to judging abnormality, and do not reflect a regurgitation of formally taught conclusions. First, our item analyses showed that clinicians' beliefs about how precipitating events should affect clinical judgments were virtually identical for both PTSD and MDD. Even though the *DSM* specifies that a traumatic life event is requisite for a PTSD diagnosis, and conversely suggests that traumatic life events of the kind used in the current study should be ignored in making MDD diagnoses, we found the same main findings across both disorders in our clinician participants. Clinicians in our study were relatively uninfluenced by *DSM* specifications about how causal life events should be used or ignored, and instead used those causal events systematically to inform their judgments about abnormality. A number of clinical theorists have pointed out that the *DSM* is internally inconsistent with respect to whether (or which) past life events should be considered in assessment (Boysen, 2007; Wakefield, 1992). Thus, the current data are also of interest in that they provide some insight into how clinicians elect to deal with this inconsistency. Across two disorders that differ distinctly with respect to prescribed use of past events in the *DSM*, clinicians themselves were completely consistent with respect to the manner in which past events influenced their judgments. Our clinician participants' use of life events to contextualize symptoms in assessing abnormality was more closely aligned with older, historical approaches to diagnosis than with the modern *DSM* system and its symptom-focused approach (Kihlstrom, 2002).

Second, the fact that experts did not focus more exclusively on the current state of the person in abnormality judgments is interesting in that the most common theoretical orientation among our participating clinicians was cognitive-behavioral (*N*=40). These findings from our study show that cognitive-behavioral and psychoanalytic clinicians were similarly influenced by precipitating events; the same basic findings were found to the same degree in cognitive-behavioral clinicians versus psychoanalytic clinicians. Overall, our findings indicate that the influence of past life events on the global assessment of

abnormality is strong, broadly applied, and resistant to specific professional training (theoretical orientation) and specific formal recommendations (*DSM*) to the contrary.

As we have described, our main goal in the current study was to examine the proportionate-response hypothesis with respect to clinicians' straightforward, direct intuitions about the global abnormality of behaviors, without including such confounding influences as consideration of the *DSM*. Again, we have previously shown in related work that abnormality judgments correlate strongly with diagnostic judgments and judgments of statistical rarity and cultural unacceptability, supporting their generalizability. A clear next step will be to examine the influence of life-event explanations on such concrete tasks as writing a clinical report; for example, does this influence exert itself when describing symptoms and mental status, when making the formal diagnosis, when rating level of functioning, or throughout the entire process?

Beyond this, our findings do not appear, in any systematic way, to be an artifact of the artificiality of the study. The current study is much less artificial than in past research; first, past work made use of artificial disorders, not *DSM* disorders as in the current study. Second, although a subset of practicing clinicians use standardized diagnostic interviews, presumably to inform their diagnoses, Rettew, Lynch, Achenbach, Dumenci & Ivanova (2009) found in a meta-analysis that diagnostic interviews only very weakly correspond to actual clinical diagnoses, and overall, they are relatively seldom used in actual practice. To the extent that this is the case, the fact that we did not incorporate diagnostic interviews into our method is unlikely to have significantly diminished the generalizability of our results. However, because systematic use of clinical interviews may be on the rise (Suppiger, In-Albon, Hendriksen, Hermann, Margraf, & Schneider, 2009), future work may nonetheless be needed to examine how life-event context influences diagnoses with and without the added factor of standardized diagnostic interview results.

Third, in prior work, the causal mechanisms explaining *how* life events had caused each client's current behaviors to come about were already explicitly provided in the case vignettes. In contrast, in real-life situations, it is much more likely that people observe others' behaviors and learn about events in their past, but in most cases would not be given the causal mechanism connecting them. Instead, people would most often have to attempt to come up with an explanation on their own, as in the current study. The current study was therefore, to our knowledge, the first to examine how the perceived cognitive difficulty of attempting to *generate* causal mechanisms between life events and disproportionate responses guides inferences about psychological abnormality.

Fourth, the vignettes, though not real-life case files, were written in the style of the *DSM* casebook and contained considerable detail. We openly solicited comments from all participants at the end of the task; our experience in the past has been that clinicians are typically quite frank if they feel that the materials are artificial or convoluted. None of the clinicians in the current study, however, said that they found the vignettes to be artificial. Two clinicians did note that the vignettes did not reflect the full spectrum of cases, such that more moderate cases were absent (i.e., in between the mildly distressed and disordered cases in terms of severity). We did this deliberately to maximize the possible degree of cause-effect mismatch in the vignettes. In line with our current arguments, however, we can predict that moderate cases would be expected to yield the same results, such that past events and current behaviors influence abnormality judgments conjointly, just as in the relatively more extreme cases in the current study. The strength of the proportionate-response effect on abnormality judgments would be dependent on the degree of mismatch between the past events and current behaviors. Additional experimentation will be needed to answer this question definitively.

Fifth, we believe that the influence of demand characteristics in our study was likely very minimal. Our key manipulation of event type *between subjects* allowed us to carry out clean paired comparisons when examining each behavior type across event types. Demand characteristics therefore cannot explain the disparity between, for example, clinicians' judgments of the unaffected behaviors in the two event conditions. Behavior type itself *was* manipulated within-subject, such that each participant saw each of the behavior types with the same event type (i.e., either all traumatic or all mildly distressing in nature). However, well-known systematic studies on within-subject designs have established that people are perfectly willing to report when manipulated information has no effect on their judgments (Baron & Hershey, 1988; Frisch, 1993). As professional clinical psychologists are acknowledged experts in the field we investigated, it seems even less likely that they would knowingly provide judgments they did not actually agree with, simply to corroborate our hypotheses or their perceptions of our hypotheses thereof.

Scope and future directions

One question that might arise is whether the proportionate-response hypothesis would be supported in clinicians' reasoning about disorders such as schizophrenia and psychotic disorders. However, even if it were found that explaining schizophrenia does not make its behavioral symptoms seem less abnormal, we argue that this would still be consistent with our claim that explanations exert a normalizing effect only to the degree that the explanations are perceived to increase understanding of the behaviors. As has previously been documented, clinicians on average believe schizophrenia to have a markedly stronger biological basis than depression (Ahn, Proctor, & Flanagan, 2009). As such, in the case of schizophrenia, knowing a precipitating life event should not be expected to lead to a strong feeling of understanding the patient, because clinicians are unlikely to perceive a life-event cause as a satisfactory explanation for symptoms that they fundamentally believe to have a significant biological underpinning. Second, as we have previously shown, so-called biological causes such as genetics or neurochemical imbalances do not lead psychiatric symptoms to be perceived as more normal than when the cause is unknown (Ahn et al., 2003). Importantly, participants in this previous study reported that they did not understand the causal mechanisms underlying the brain-based cases nearly as well as they understood those underlying the life-event cases. All of this is compatible with our main claim that judgments of abnormality are influenced by explanations *to the degree* that those explanations lead to the perception of better understanding the symptoms. Until people acquire a better everyday understanding and working knowledge of the mechanisms by which biological causes give rise to disorders, we can expect the normalizing effect to be generally weaker for disorders perceived to have strong biological causes. Relatedly, we might expect brain experts such as neurologists, psychiatrists, and neurobiologically-oriented psychologists to be influenced by biological causes in the same way as shown here for life-event causes. That is, the more readily a person is able to explain behaviors given the (biological) cause, given his or her background knowledge, the more normal the client may appear, and vice-versa. It also remains for future research to determine whether our findings would also extend to biological or medical symptoms (as opposed to psychological or behavioral ones), and to problems or malfunctions in other domains of study.

One boundary condition on the proportionate-response hypothesis is that there is likely to be a limited time frame within which we can reasonably expect it to apply. Consider again the hypothetical case vignette of Justin, who experienced atrocities in military combat, and felt absolutely contented and peaceful afterwards. For this specific vignette (see supplementary appendix for the full text), the clinicians in our study judged that having absolutely no negative response to the traumatic experience of atrocities in combat was just as abnormal as having full-blown PTSD symptoms (p=.8), just as we found for the full data set. Intuitively,

this finding makes sense; it seems reasonable to expect distress of at least some duration in the subset of military personnel who are directly exposed to traumatic incidents or atrocities. For example, studies with Vietnam veterans indicate a strong correlation between exposure to atrocities in combat and the subsequent severity of PTSD symptoms (e.g., Beckham, Feldman, & Kirby, 1998). However, although clinicians might normally expect a certain level of distress in response to trauma in combat, they might also reasonably expect that the distress should ideally become attenuated over time. That is, if the distress continues to persist at intense levels for many years, this may no longer be considered to be a proportionate response. Clinicians' intuitions regarding exactly how long the time window should reasonably extend to produce proportionate-response effects are likely to differ somewhat for different disorders, and remain an important issue open to systematic investigation.

It will also be important in future work to examine reasoning about cases in which there is a confluence of events over time that lead to depressive behaviors, as opposed to a single clear, precipitating event (Hammen, 1992). Our prediction, drawn directly from the *proportionate-response hypothesis*, is that the same effect on judgments of abnormality and other clinical judgments would be obtained, to the degree that such a confluence of events increases perceived understanding of the behaviors. In other words, our proportionate-response claim is not tied to the requirement that the symptoms be precipitated by a single life event. It is possible, of course, that people will find it more difficult to understand behaviors in response to a confluence of many life events than to a single event.

Finally, Schön's (1983) classic analysis of practitioners' thinking suggests a related avenue of potential future inquiry. Schön argued that practitioners in action are primarily motivated by the goal of eliciting change, and that their perceptions of increased understanding of their clients are often incidental to observing the results of their manipulations to effect change. Whenever change is indeed elicited by a therapeutic manipulation, according to Schön, the practitioner will see this as evidence supporting his or her hypothesis regarding what had caused the undesirable state. For example, after interviewing a client, a clinician may hypothesize that the client's symptoms stem from a traumatic combat experience, and may test that hypothesis by directly addressing that combat experience in therapy. If a particular explanatory hypothesis seems unlikely to the clinician (e.g., because of a proportionate-response mismatch), the clinician may be very unlikely to ever test it with an intervention. If, on the other hand, there is a proportionate-response match, the clinician is very likely to test it via therapeutic intervention and, in the process, ultimately self-reinforce the perception of understanding the client. These possibilities await further research.

Implications of the current findings for clinical practice

Although we have not thus far presented the current findings as reflective of the use of a heuristic, it is possible that clinicians might apply something along the lines of a proportionate-response "heuristic," particularly in cases in which there is uncertainty in the judgment (e.g., clinicians are relatively unfamiliar with the presenting symptoms; the symptoms do not map to a clear diagnosis). As has been extensively argued in the classic judgment and decision making literature, the use of heuristics in general tends to carry the benefit of enabling fast and frequently accurate decision making, yet typically also has localized costs, usually in that it leads to inaccurate judgments in a minority of cases (Arkes, 1991; Tversky & Kahneman, 1974). Costs notwithstanding, it is possible for decision-making strategies based on simple heuristics to allow reasoners to outperform more complex algorithms (Gigerenzer & Goldstein, 1996).

Thus, an important practical question remains as to whether it is in fact desirable or undesirable for clinicians to allow past events to influence assessments of current behaviors

in the manner observed in the current study. Thus far, we have argued that our findings reflect clinicians' true, untrained beliefs about psychological abnormality. Experienced clinicians, who ostensibly knew perfectly well what the *DSM* recommends, factored past events into their assessment decisions, and informal follow-up comments volunteered by a number of the clinicians in our study suggest that this effect was, in many cases, the result of deliberate decision-making. Indeed, previous work has shown that clinical psychologists across theoretical orientations formulate their own idiosyncratic theories about the etiology of disorders (Kim & Ahn, 2002), even if the *DSM* does not describe or endorse them. It has also been argued elsewhere that highly effective mental health clinicians across theoretical orientations often either deliberately or implicitly place emphasis on the past and how it gave rise to current behaviors (Ablon & Jones, 1998; Shedler, 2010).

Exactly how clinicians' intuitions may interfere with the direct application of *DSM* guidelines must be examined systematically in additional work; however, the sociological literature indicates that the overall outcome is likely somewhere in between, such that the *DSM* clearly has a strong influence, but clinicians' own opinions give rise to fudging and workarounds when the two do not agree (Whooley, 2010). To the degree that the *DSM* does guide diagnostic decisions, there is considerable disagreement in the field about the potential consequences of having a diagnostic system that directs clinicians to sidestep questions of context for the symptoms they observe in clients. Some critics of the *DSM* symptom-based approach to the diagnosis of Major Depressive Disorder, for example, have argued that ignoring contextual factors and questions of etiology has resulted in an over-pathologizing of what society has historically considered to be relatively "normal" distress in response to stressful events and situations. One argument is that this, in turn, may have contributed to an over-prescription of antidepressant therapy, which may be less effective or even ineffective for moderate to milder cases of MDD (Horowitz & Wakefield, 2007). This critique has been extended to the context of chronic illness, wherein it is unclear whether the high rates of depression that are observed in patients with diabetes, for example, are indicative of comorbidity between MDD and diabetes or reflective of reasonable distress caused by the burden of illness. This distinction is much more than a question of semantics, and is likely to have important implications for the selection of appropriate approaches to treatment (Gonzalez, Fisher, & Polonsky, 2011).

We speculate that in clinical practice, clinicians may, in effect, be partially compensating for the above issues. As we have discussed, allowing past events to affect assessments of current behaviors runs contrary, in many ways, to formal *DSM* recommendations for most specific disorders. However, given that the *DSM* was created by a large set of widely acknowledged experts in the field, is disseminated in a number of countries, serves as the core basis for communication among researchers and clinicians, and at least in the U.S., is considered the authoritative basis for diagnoses by health insurance companies, there are likely to be practical and real-world consequences for not adhering to the *DSM*, as well (Ahn & Kim, 2008). For one, allowing past events to explain away disordered behaviors could potentially result in under-diagnosis and perhaps even ultimately undertreatment of real cases of mental disorders (Maj, 2008). One could also conversely imagine over-diagnosis and overtreatment of cases in which individuals are not currently displaying distressed or disordered behaviors after experiencing significant life stressors. Although such stressors could potentially generate future distress in such cases, it may nonetheless be counterproductive to label unperturbed behaviors as pathological before the fact. Overall, given the current data, we recommend that clinicians operate with the awareness that causal context may influence their global judgments about the abnormality of a client's behaviors.

Acknowledgments

References

Ablon JS, Jones EE. How expert clinicians' prototypes of an ideal treatment correlate with outcome in psychodynamic and cognitive-behavioral therapy. Psychotherapy Research. 1998; 8:71–83.

Ahn W, Kim NS. Causal theories of mental disorder concepts. Psychological Science Agenda. 2008; 22:3–8.

Ahn W, Novick L, Kim NS. "Understanding it makes it normal:" Causal explanations influence person perception. Psychonomic Bulletin & Review. 2003; 10:746–752. [PubMed: 14620373]

Ahn W, Proctor CC, Flanagan EH. Mental health clinicians' beliefs about the biological, psychological, and environmental bases of mental disorders. Cognitive Science. 2009; 33:147–182. [PubMed: 20411158]

American Psychiatric Association. Diagnostic and statistical manual of mental disorders (2nd Ed.). Washington, DC: Author; 1968.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders (3rd Ed.). Washington, DC: Author; 1980.

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

American Psychiatric Association. Diagnostic and statistical manual of mental disorders (4th Ed., text revision). Washington, DC: Author; 2000.

American Psychological Association. The directory of ethnic minority professionals in psychology (4th Ed.). Washington, DC: Author; 2001.

Arkes HR. Costs and benefits of judgment errors: Implications for debiasing. Psychological Bulletin. 1991; 110:486–498.

Baron J, Hershey JC. Outcome bias in decision evaluation. Journal of Personality and Social Psychology. 1988; 54:569–579. [PubMed: 3367280]

Beck, JS. Cognitive therapy: Basics and beyond. New York: Guilford Press; 1995.

Boysen GA. An evaluation of the *DSM* concept of mental disorder. The Journal of Mind and Behavior. 2007; 28:157–174.

Carson RC. Aristotle, Galileo, and the DSM taxonomy: The case of schizophrenia. Journal of Consulting and Clinical Psychology. 1996; 64:1133–1139. [PubMed: 8991300]

Clark, DA.; Beck, AT. Scientific foundations of cognitive theory and therapy of depression. New York: John Wiley & Sons; 1999.

de Kwaadsteniet L, Hagmayer Y, Krol NPCM, Witteman CLM. Causal client models in selecting effective interventions: A cognitive mapping study. Psychological Assessment. 2010; 22:581–592. [PubMed: 20822270]

Einhorn HJ, Hogarth RM. Judging probable cause. Psychological Bulletin. 1986; 99:3–19.

Frisch D. Reasons for framing effects. Organizational Behavior and Human Decision Processes. 1993; 54:399–429.

Gigerenzer G, Goldstein DG. Reasoning the fast and frugal way: models of bounded rationality. Psychological Review. 1996; 103:650–669. [PubMed: 8888650]

Gold SD, Marx BP, Soler-Baillo JM, Sloan DM. Is life stress more traumatic than traumatic stress? Anxiety Disorders. 2005; 19:687–698.

Gonzalez JS, Fisher L, Polonsky WH. Depression in diabetes: Have we been missing something important? Diabetes Care. 2011; 34:236–239. [PubMed: 21193621]

Hammen C. Cognitive, life stress, and interpersonal approaches to a developmental psychopathology model of depression. Development and Psychopathology. 1992; 4:189–206.

Hatala R, Norman GR, Brooks LR. Impact of a clinical scenario on accuracy of electrocardiogram interpretation. Journal of General Internal Medicine. 1999; 14:126–129. [PubMed: 10051784]

Horwitz, AV.; Wakefield, JC. The loss of sadness: How psychiatry transformed normal sorrow into depressive disorder. New York, NY: Oxford University Press; 2007.

Kahneman, D.; Tversky, A. The simulation heuristic. In: Kahneman, D.; Slovic, P.; Tversky, A., editors. Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press; 1982. p. 201-208.

Kessler RC, Neighbors HW. A new perspective on the relationships among race, social class, and psychological distress. Journal of Health and Social Behavior. 1986; 27:107–115. [PubMed: 3734380]

Kihlstrom, JF. To honor Kraepelin…: From symptoms to pathology in the diagnosis of mental illness. In: Beutler, LE.; Malik, ML., editors. Rethinking the DSM: A psychological perspective. Washington, DC: American Psychological Association; 2002.

Kim NS, Ahn W. Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. Journal of Experimental Psychology. 2002; 131:451–476. [PubMed: 12500858]

Kim NS, LoSavio ST. Causal explanations affect judgments of the need for psychological treatment. Judgment and Decision Making. 2009; 4:82–91.

Kim, NS.; Nguyen, TP.; Paulus, DJ. Explanatory context overrides formal category definitions in clinical experts' diagnostic judgments. In: Carlson, L.; Hölscher, C.; Shipley, T., editors. Proceedings of the Thirty-Third Annual Conference of the Cognitive Science Society; Cognitive Science Society; Austin, TX; 2011.

Lehman AK. The content and organization of therapists' mental representations of their patients and the psychotherapy process (Doctoral dissertation, Yale University, 1991). Dissertation Abstracts International. 1992; 53:1067.

Lifton, RJ. Understanding the traumatized self. In: Wilson, JP.; Harel, Z.; Kahana, B., editors. Human adaptation to extreme stress: From the Holocaust to Vietnam. New York: Plenum Press; 1988. p. 7-31.

Lincoln A. Psychiatric emergency room decision-making, social control and the "undeserving sick.". Sociology of Health and Illness. 2006; 28:54–75. [PubMed: 16509942]

Maj M. Depression, bereavement, and "understandable" intense sadness: Should the *DSM-IV* approach be revised? American Journal of Psychiatry. 2008; 165:1373–1375. [PubMed: 18981066]

Medin DL. Concepts and conceptual structure. American Psychologist. 1989; 32:1469–1481. [PubMed: 2690699]

Meehl, PE. Psychodiagnosis: Selected papers. Minneapolis, MN: University of Minnesota Press; 1973. Why I do not attend case conferences; p. 225-302.

Norman GR, Arfai B, Gupta A, Brooks LR, Eva KW. The privileged status of prestigious terminology: Impact of "medicalese" on clinical judgments. Academic Medicine. 2003; 78:S82–S84. [PubMed: 14557104]

Pennington N, Hastie R. Explanation-based decision making: Effects of memory structure on judgment. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1988; 14:521–533.

Rettew DC, Lynch A, Achenbach TM, Dumenci L, Ivanova MY. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. International Journal of Methods in Psychiatric Research. 2009; 18:169–184. [PubMed: 19701924]

Rosenfield S. Sex roles and societal reactions to mental illness: The labeling of "deviant" deviance. Journal of Health and Social Behavior. 1982; 23:18–24. [PubMed: 7077067]

Schön, DA. The reflective practitioner: How professionals think in action. New York: Basic Books; 1983.

Schwartz S, Link BG. The "well control" artefact in case/control studies of specific psychiatric disorders. Psychological Medicine. 1989; 19:737–742. [PubMed: 2798641]

Seides R. Should the current DSM-IV-TR definition for PTSD be expanded to include serial and multiple microtraumas as aetiologies? Journal of Psychiatric and Mental Health Nursing. 2010; 17:725–731. [PubMed: 21050339]

Shedler J. The efficacy of psychodynamic psychotherapy. American Psychologist. 2010; 65:98–109. [PubMed: 20141265]

Smith J. Reexamining psychotherapeutic action through the lens of trauma. Journal of the American Academy of Psychoanalysis. 2004; 32:613–631.

Spitzer, RL.; Gibbon, M.; Skodol, AE.; Williams, JBW.; First, MB., editors. DSM-IV-TR casebook. Washington, DC: American Psychiatric Press; 2001.

Suppiger A, In-Albon T, Hendriksen S, Hermann E, Margraf J, Schneider S. Acceptance of structured diagnostic interviews for mental disorders in clinical practice and research settings. Behavior Therapy. 2009; 40:272–279. [PubMed: 19647528]

Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. Science. 1974; 185:1124–1131. [PubMed: 17835457]

Wakefield JC. Disorders as harmful dysfunction: A conceptual critique of the *DSM-III-R*'s definition of mental disorder. Psychological Review. 1992; 47:373–388.

Wakefield JC. Diagnosing *DSM-IV*—Part I: *DSM-IV* and the concept of disorder. Behaviour Research and Therapy. 1997; 35:633–649. [PubMed: 9193127]

Whooley O. Diagnostic ambivalence: Psychiatric workarounds and the *DSM*. Sociology of Health & Illness. 2010; 32:452–469. [PubMed: 20415790]

Yopchick, JE.; Kim, NS. Causal explanations guide inferences about children's bizarre behaviors. In: Carlson, L.; Hölscher, C.; Shipley, T., editors. Proceedings of the Thirty-Third Annual Conference of the Cognitive Science Society; Cognitive Science Society; Austin, TX. 2011.

Yopchick JE, Kim NS. The influence of causal information on judgments of treatment efficacy. Memory & Cognition. 2009; 37:29–41.

Young ME, Brooks LR, Norman GR. Found in translation: The impact of familiar symptom descriptions on diagnosis in novices. Medical Education. 2007; 41:1146–1151. [PubMed: 18045366]

Zalaquett CP, Fuerth KM, Stein C, Ivey AE, Ivey MB. Reframing the DSM-IV-TR from a multicultural / social justice perspective. Journal of Counseling and Development. 2008; 86:364–371.
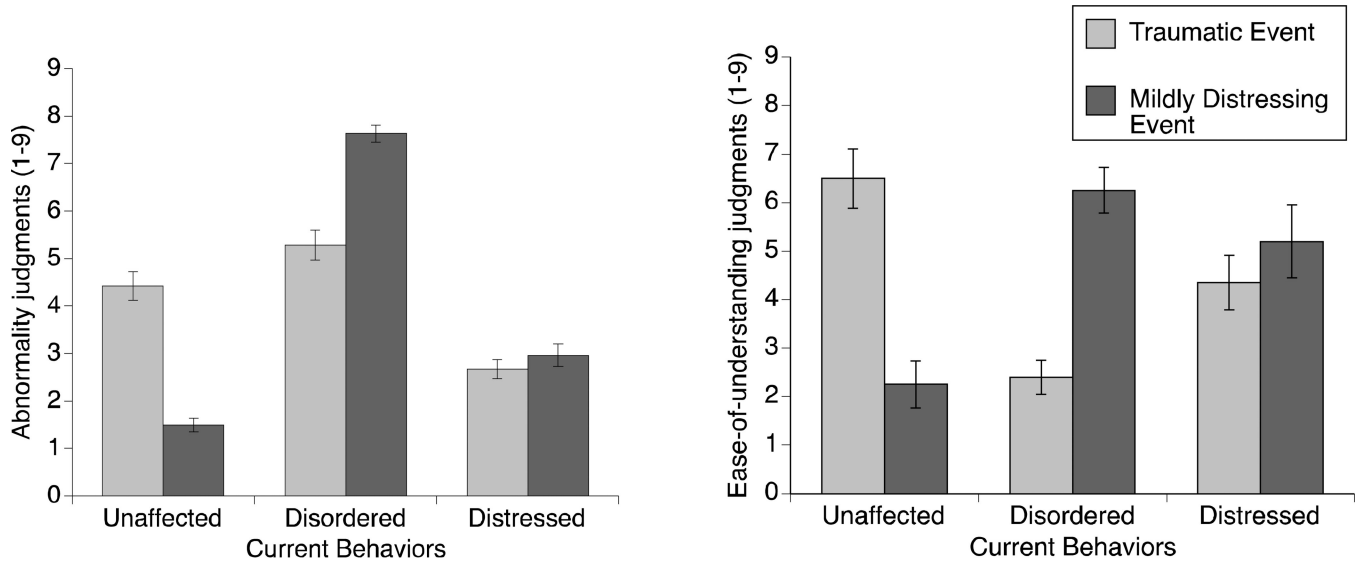
(A) Abnormality judgments   (B) Ease-of-understanding judgments

Figure 1.
Clinicians' mean abnormality and ease-of-understanding judgments. (A) Higher ratings indicate greater perceived abnormality of behaviors. (B) Higher ratings indicate greater difficulty in understanding the behaviors. Error bars indicate standard errors.
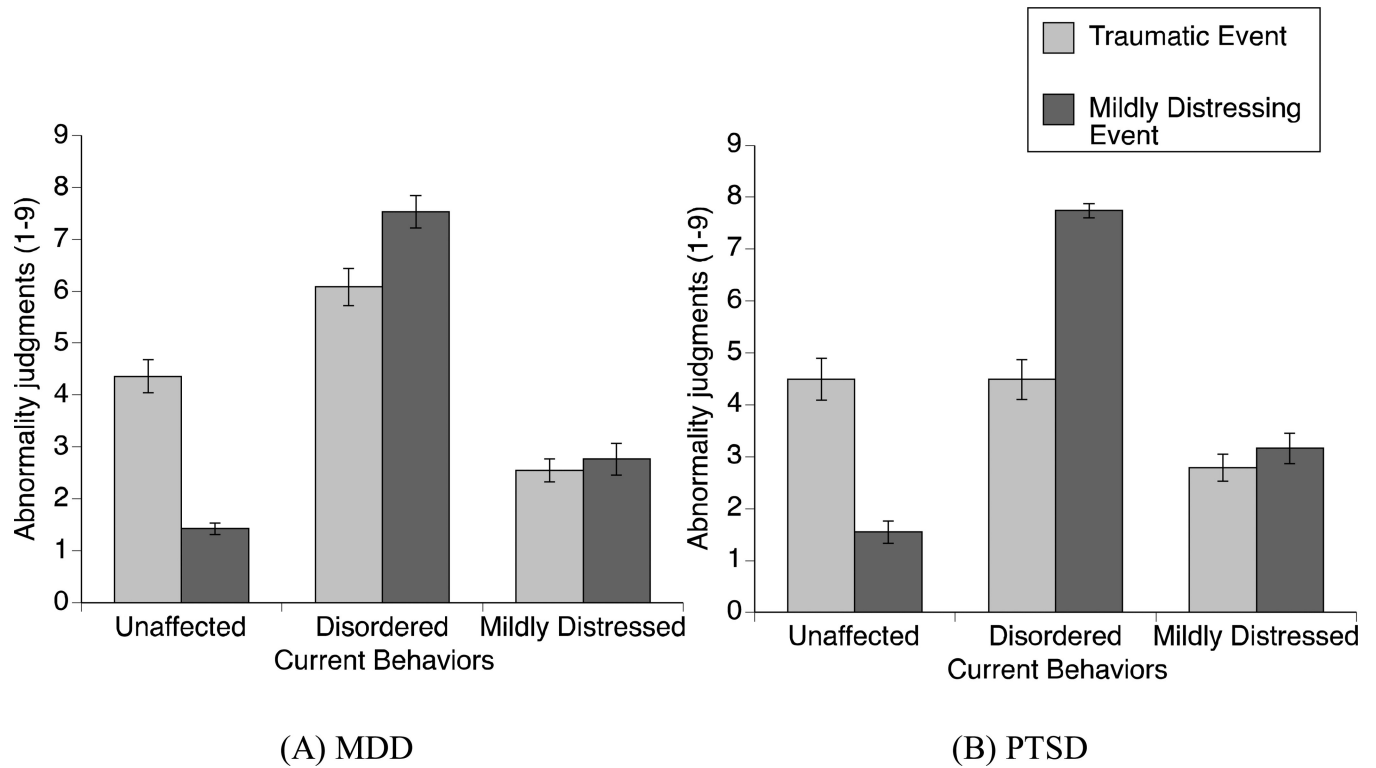
(A) MDD   (B) PTSD

Figure 2.
Clinicians' mean abnormality ratings, broken down by symptom type: (A) MDD and (B) PTSD. Higher ratings indicate greater perceived abnormality of behaviors. Error bars indicate standard errors.