



Published in final edited form as:

Psychol Bull. 2006 May ; 132(3): 477–495.

Inventorying Stressful Life Events as Risk Factors for Psychopathology: Toward Resolution of the Problem of Intracategory Variability

Bruce P. Dohrenwend

Columbia University and the New York State Psychiatric Institute

Abstract

An explosion of research on life events has occurred since the publication of the Holmes and Rahe checklist in 1967. Despite criticism, especially of their use in research on psychopathology, such economical inventories have remained dominant. Most of the problems of reliability and validity with traditional inventories can be traced to the intracategory variability of actual events reported in their broad checklist categories. The purposes of this review are, first, to examine how this problem has been addressed within the tradition of economical checklist approaches; second, to determine how it has been dealt with by far less widely used and far less economical labor-intensive interview and narrative-rating approaches; and, third, to assess the prospects for relatively economical, as well as reliable and valid, solutions.

Keywords

stressful life events; checklist inventories; intracategory variability; narrative-rating methods; psychopathology

Almost four decades ago, Holmes and Rahe (1967) published a checklist of 43 events, such as *death of a spouse, divorce, fired at work, and sex difficulties*, called the Schedule of Recent Experiences (SRE). Its purpose was to inventory “fundamentally important environmental incidents” (Meyer, 1951, p. 53) that were found (in analyses of patients’ charts) to frequently precede illness onsets. Stressful events were defined as occurrences that were likely to bring about readjustment-requiring changes in people’s usual activities. A magnitude estimation procedure was used by panels of judges to assign Life Change Unit (LCU) scores to each of the 43 events on the list (Holmes & Rahe, 1967). A summation of these scores on events occurring in a given, usually quite recent, period of time was taken as the indicator of amount of stress. The checklist can be answered in either a self-administered questionnaire or an interview.

Since the publication of this economical measurement procedure, the SRE, a tremendous increase has occurred in the construction of such measures and in quantitative research on relations between inventoried life events and health. For example, a search of the terms *life events, life change, stressful life events, and life stress* (or a combination of these terms) using PsycINFO (<http://www.apa.org/psycinfo>) shows an increasing rate of publications on these topics, from 292 in the decade of 1967 to 1976, to 2,126 in 1977 to 1986, to 4,269 in 1987 to

Correspondence concerning this article should be addressed to Bruce P. Dohrenwend, Social Psychiatry Research Unit, Columbia University, 100 Haven Avenue, Tower 3–19E, New York, NY 10032. E-mail: dohrenw@pi.cpmc.columbia.edu.

Bruce P. Dohrenwend, Department of Psychiatry and Mailman School of Public Health, Columbia University; New York State Psychiatric Institute.

1996, to 3,341 in the truncated 1997 to 2005 portion of the present decade. This voluminous literature documents that life events are related to a wide variety of physical and psychological problems in both cross-sectional and longitudinal research (e.g., Breslau, 2002; Brown & Harris, 1989; Dohrenwend & Dohrenwend, 1974; Grant, Compas, Thurm, McMahon, & Gipson, 2004; Gunderson & Rahe, 1974; Paykel, 1974; Rahe & Arthur, 1978).

Most of the investigators in this research have used the Holmes and Rahe (1967) instrument, other, usually longer, inventories (e.g., Dohrenwend, Krasnoff, Askenasy, & Dohrenwend, 1978; Paykel, Prusoff, & Uhlenhuth, 1971; Sarason, Johnson, & Siegel, 1978), or unnamed admixtures of several instruments (e.g., Steele, Henderson, & Duncan-Jones, 1980) to investigate relations between stress and health outcomes. Special inventories have been developed for subgroups in the population. For example, at least 16 different checklists have been used in research with children and adolescents (Grant et al., 2004), and additions have been made to existing scales to take into account events in special circumstances, such as experiences with terrorism during the continuing hostilities in which Israel has been involved (Levav, Krasnoff, & Dohrenwend, 1981). Moreover, inventories have been developed specifically for the study of extreme situations, such as combat in the first Gulf War (Southwick, Morgan, Nicolau, & Charney, 1997) and the plight of Cambodian refugees (Mollica, Poole, & Tor, 1998). The common characteristic of these traditional checklists for research, whether focusing on usual situations or extreme situations, is that they consist of rather broad categories of events (e.g., divorce) rather than detailed descriptions of individual events (e.g., a divorce after a period of marital conflict over the infidelity of one's spouse).

Although the emphasis has usually been on recent events in relatively brief intervals of time, such as 6 months to 1 year, a few studies have attempted to assess "traumatic" events (usually life-threatening or otherwise threatening to physical integrity), other major events (e.g., spousal bereavement), or both over the life course of the respondents (Breslau, Davis, Andreski, & Peterson, 1991; Breslau, Davis, Peterson, & Schultz, 1997; Breslau et al., 1998; Davidson, Hughes, Blazer, & George, 1991; Kessler, Davis, & Kendler, 1997; Kessler, Sonnega, Bromet, & Nelson, 1995; Norris, 1992; Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993; Turner & Lloyd, 1995, 2004). These attempts at lifetime coverage of major events are consistent with Meyer's conception of a *life chart of fundamentally important environmental incidents* and have the potential of providing more comprehensive information about (a) whether the events recur or persist over time to the extent of becoming ongoing difficulties (Brown & Harris, 1978) or chronic stressors (e.g., Lepore, 1997); (b) the role of the events singly and in relation to earlier stressful events in both onset and course of disorder; and (c) reciprocal relations between the occurrence of some types of events that are not completely independent of the behavior of the individual and the occurrence or recurrence of episodes of disorder (Grant et al., 2003, 2004; Hammen, 2005; Monroe & Harkness, 2005). When lifetime coverage involves long periods of recall, as it does in all of the previously mentioned retrospective studies, it is only feasible to focus on major events. Assessment of more easily forgotten minor events is not even remotely feasible unless recall periods are as short as possible. For example, longitudinal questioning at monthly intervals for 10 months has been found to elicit an average of 16.7 checklist events, most of them minor, compared with an average of 6.0 major and minor events reported retrospectively for the same 10-month period (Raphael, Cloitre, & Dohrenwend, 1991).

Although most attempts to inventory stressful life events have been conducted with traditional checklists of relatively broad life event categories such as those in the SRE, some more recent investigators have used checklists with more narrowly defined event categories to investigate both relatively large events (e.g., Paykel, 1997; Kubany et al., 2000) and small daily incidents or "hassles" (e.g., Kanner, Coyne, Schaefer, & Lazarus, 1981; Zautra, Guarnaccia, & Dohrenwend, 1986). The most radical methodological departure from traditional checklists,

however, has been provided by much less economical procedures that involve intensive semistructured interview and rating procedures. The interviews are designed to elicit details of what actually happened to provide narratives of the events that can be rated by trained investigators for severity and other important characteristics. The best known and most widely used of these procedures was developed by George Brown and colleagues and is called the Life Events and Difficulties Schedule (LEDS; Brown & Harris, 1978). In contrast to traditional checklists of broad event categories and more recent checklists of more narrowly defined categories, LEDS and other similarly labor-intensive procedures (e.g., Dohrenwend, Raphael, Schwartz, Stueve, & Skodol, 1993; Hammen, 1991) will be referred to as narrative-rating approaches.

The emphasis in this paper is on stressful life events in relation to psychopathology. Much of the methodological analysis is also relevant for physical health outcomes. However, some problems, such as the confounding of the measurement of events and the measurement of health outcomes, are more serious when the outcomes involve psychiatric disorders for which no laboratory tests exist that are independent of self-reported symptoms. In addition, there may also be differences in the characteristics of events that are important for physical health outcomes versus outcomes involving psychopathology. Further investigation of these important differences is beyond the scope of this paper.

Checklists have continued to be dominant in research on the role of stressful life events in psychopathology, as well as other health outcomes. All of the previously discussed studies of potentially traumatic and other major negative events focused on various types of psychopathology and used traditional checklists with broad-event categories to inventory events over the life course. Checklists in the form of self-administered questionnaires or in the form of structured interviews consisting of closed questions with fixed alternative response categories have been relied on in other influential research on the role of life events in psychopathology conducted in recent years. The behavioral genetics study of depression by Kendler et al. (1995), the nationwide epidemiological research on the comorbidity of psychiatric disorders by Kessler et al. (1995), and, most recently, the study of gene-by-environment interaction for depression by Caspi et al. (2003). are examples.

The vast body of research on life events and psychopathology that has relied on checklist inventories has been subject to many critical reviews (see reviews in Brown, Sklair, Harris, & Birley, 1973; Cohen, Kessler, & Gordon, 1997a; Day, 1989; Dohrenwend, 2000; Gorman, 1993; Grant et al., 2003; Kessler, 1997; Lloyd, 1980; Mechanic, 1975; Monroe & Roberts, 1990; Norman & Malla, 1993; Paykel, 2001; Rabkin & Struening, 1976; Rutter, 1986; Tennant, Bebbington, & Hurry, 1981; Thoits, 1995; Turner & Wheaton, 1997). Although the reviewers have been impressed with the consistent relationships found between life events and psychological problems using such procedures, two broad themes of criticism can be found in these reviews. One theme emphasizes the small to modest amount of variance explained in relationships of the events to psychopathology and points to other variables, such social support, that may increase or decrease the impact of the events (e.g., Johnson & Bradlyn, 1988; Paykel, 2001; Rabkin & Struening, 1976; Sarason & Sarason, 1985). The other theme of criticism focuses more on problems of reliability and validity evident in the traditional checklist measures of stressful life events used in most of this research (e.g., Brown, Sklair, Harris & Birley, 1973; Creed, 1993; Dohrenwend, 1974, 2000). These two themes are not, of course, mutually exclusive. Focus here, however, is more on the latter, that is, on problems in the conceptualization and measurement of stressful life events that make the results of the research difficult to interpret.

The fundamental methodological puzzle in inventorying life events as risk factors for psychopathology is how to solve the problem of intracategory variability in traditional checklist

measures—that is, the fact that a variety of types of experience are encompassed by each particular event category (Dohrenwend et al., 1993). Although various investigators have addressed one or another of the consequences of intracategory variability (e.g., Brown, 1974; Dohrenwend et al., 1993; Katschnig, 1986; McQuaid et al., 1992), they have done so piecemeal. No systematic review of the nature of the problem, its full range of sequelae, and the adequacy of various solutions has been conducted. The purposes here are, first, to examine the nature of the problem of intracategory variability and its full range of consequences; second, to describe how some of its consequences have been addressed within the tradition of economical checklist procedures and by narrative-rating procedures; and, third, to assess both the problems and prospects of the checklist procedures, on the one hand, and narrative-rating procedures, on the other, for reducing intracategory variability.

The problems with, and prospects of, checklist approaches are considered first, followed by a similar analysis of the more labor-intensive and time-consuming narrative-rating procedures. These analyses are followed by an examination of the handful of studies that have attempted to compare checklist and narrative-rating measures. Finally, ideas about future research that could resolve the problem of intracategory variability and result in the development of practical, as well as reliable and valid, procedures for inventorying stressful life events as risk factors for psychopathology are proposed.

Problem of Intracategory Variability

Traditional checklist approaches describe events at such a general level—for example, *marriage, divorce, death of a spouse* (Holmes & Rahe, 1967)—that they might more accurately be called *lists of topics* or *categories of events* than *events* per se. The core problem with traditional checklists, and the one to which all of the others are in some way and to some degree related, is that the actual experiences that lead a respondent to make a positive response to a given checklist category vary greatly. Findings from a study of 429 adult residents of the Washington Heights section of New York City illustrate this problem (Dohrenwend, Link, Kern, Shrout, & Markowitz, 1990). These respondents were interviewed about their responses to the 102 categories on a previously developed traditional checklist called the Psychiatric Epidemiology Research Interview (PERI) Life Events Scale (Dohrenwend et al., 1978).

When positive responses to event categories were probed for details of what actually occurred, it became evident, for example, that some *deaths of close friends* turned out to involve deaths of long absent, childhood friends to whom the respondents were no longer close; that *serious illness and injury* events ranged from episodes of flu and sprained arms to severe heart attacks; and that *laid off*, intended to encompass economic failure of the employer, was sometimes a euphemism for being fired for cause. Clearly, a positive response for the same item on the checklist could, and did, in fact, represent very different types of actual experience.

To investigate this matter further, the subsequent descriptions of what actually happened that led to a positive response to a checklist item by respondents were rated by the investigators on such characteristics as the likely magnitude of change in usual activities that the event would induce in most people who experience it.

As can be seen in Table 1, the anchoring examples used by the raters present experiences ranging from *the catastrophic* to *the trivial*.

Table 2 shows, for 10 checklist items, the ratings by the research team of likely magnitude of change based on respondents' accounts of what actually happened. Each of the 10 items met the following criteria: were reported present by more than 20 respondents; were rated as likely to be negative or undesirable for most people who experience this type of event (Dohrenwend et al., 1978); and had one of the 10 largest a priori Life Change Unit (LCU) scores (Holmes &

Rahe, 1967) among the checklist items meeting the preceding criteria. To arrive at the LCU scores for the PERI event categories, the magnitude estimation procedure used by Holmes and Rahe (1967) was followed. This involved asking judges to rate the magnitude of change for each item compared with *marriage*, which was given an arbitrary score of 500. The resulting LCU scores for the PERI event categories ranged from 1,036 for *child died*, which was ranked 1, to 163 for *acquired a pet*. Marriage, the modular event with 500, was ranked 22.

As can be seen in Table 2, a wide variability in narrative ratings exists within most of the event categories. Moreover, even when a majority or near majority of event descriptions in a particular category receive the same rating on amount of change, as in *death of relative other than spouse or child* and *death of close friend*, the findings are surprising. The modal ratings of the narratives describing the events in these categories were *no change lasting a week or more*, despite the fact that these items ranked in the top quarter of events on the PERI checklist in terms of the magnitude of their LCU scores. The actual descriptions of what happened for the events reported in these categories show why the discrepancy occurred. Most of the deaths were of biologically distant relatives who were living far away at the time and of acquaintances or long-absent friends with whom the respondents were no longer close.

Note also the category *laid off work*. If being *laid off* was the consequence of an event like a plant shutdown, then it is difficult to see how that experience could lead to little or no change for almost a quarter of the respondents with such an experience. Again, the respondents' descriptions showed what had happened. Some of those *laid off*, for example, had been employed as musicians, dancers, or actors who had come to the end of a contractual engagement. For them, this was hardly an event at all, because it was an expected and normal part of their occupational lives to be *laid off* several times a year. This heterogeneity with regard to the magnitude of events grouped into a particular checklist category is important because it obscures important differences, such as the paramount role of major negative events in first episodes of major depression and the paramount role of minor negative events in episodes that represent frequent recurrences of major depression (see updated review by Monroe & Harkness, 2005).

The problem of intracategory variability is evident in other critical characteristics of events as well—characteristics such as their valence and source (Dohrenwend, 2000). For example, being *laid off* on this checklist was meant to indicate the negative event of loss of a job because of the fateful, external environmental factor of economic failure by the respondent's employer. When positive responses to this item were probed for details of what actually happened, it became clear that sometimes *laid off* was a euphemism for being fired for cause, an experience that cannot be categorized as fateful in the previously discussed sense (Dohrenwend, 1979). For 30.7% of the respondents who were *laid off*, the sequence leading to the event was rated as being equally or mostly determined by the respondents' behavior, as opposed to external circumstances; for 43.3% of the respondents, this supposedly negative event was rated as *mixed positive and negative*; and for another 3.3% as *mainly positive*.

Summary of Problem of Intracategory Variability

The core problem with traditional checklist inventories is that the actual responses to the broad categories of events on the lists vary greatly. As a result, minor illnesses are included along with major events such as heart attacks in response to an item like *serious physical illness or injury*. Overinclusiveness with regard to magnitude is important because it is major events, not minor ones, which have proved to be substantial risk factors for some types of psychopathology, such as first episodes of major depression. The problem of intracategory variability is not limited to magnitude. It extends to other fundamentally important characteristics of events, for example, their valence (i.e., Are they mainly positive or negative? Do they involve gain rather than loss? Are they desirable rather than undesirable?) and their

source (i.e., Is the origin of the events in environmental circumstances rather than a consequence of the actions of the individual?).

Impact of Intracategory Variability on Reliability and Validity of Traditional Checklist Approaches

Intracategory variability underlies more usually recognized problems of reliability and validity in traditional checklists: four sequelae of intracategory variability are unreliability of recall, susceptibility to recall bias, lack of criterion validity, and problems with construct validity.

Unreliability of Recall—Very few studies have been designed explicitly to test the reliability of traditional checklists. One of the best was conducted by Steele et al. (1980) with a checklist of items adapted from the work of Holmes and Rahe (1967) and Paykel (1974). The 52 respondents were recruited from diverse medical services. The Time 1 test consisted of a self-administered checklist questionnaire inquiring about events during the preceding 12 months; the questionnaire was followed immediately by an interview in which more detail was obtained about the Time 1 events reported. The retest 7 to 14 days later consisted of the same sequence of questionnaire followed by interview. The investigators report a test-retest correlation of .94 for total number of events. Even in this study with its very short test-retest interval, however, there was no more than 70% agreement for all of the events reported in the same category on either test. Steele et al. note that details obtained in the interviews showed that some of the discrepancies were the result of the same event being reported under different checklist categories; unfortunately the investigators do not provide information about how large a problem this was.

When the recall period is longer, reliability drops precipitously for both total scores and individual events. Reliability is particularly poor when the checklist is self-administered (see reviews by Neugebauer, 1983; Paykel, 1983; Rabkin & Struening, 1976; Thoits, 1983). For example, Neugebauer (1984) found test-retest correlations for total scores usually to be between .30 and .60. In the studies he reviewed, the investigation of reliability was sometimes a by-product of longitudinal research in which there was an overlap in reporting period covered in the first and second waves of data collection that provided two sets of reports for events occurring in the same period of time.

Problems of recall are particularly evident in comparisons of longitudinal and cross-sectional studies, which show that far more events are reported longitudinally than retrospectively. In such studies, retrospective reports are compared with longitudinal reports at shorter intervals either for the same period of time (Klein & Rubovits, 1987; Raphael et al., 1991) or a similar but nonoverlapping period of time (Monroe, 1982). In the study by Raphael et al. (1991), for example, there was longitudinal/retrospective agreement on only 25% of the events, with far more events reported in the monthly longitudinal assessment than in the retrospective assessment of the same 10-month period, as was noted earlier.

Susceptibility to Recall Bias—Checklists are rarely administered in daily diaries, and when they are, coverage tends to be limited to a few weeks or months in the life of the respondents (Eckenrode & Bolger, 1997). More usually, a substantial period of recall is involved between the occurrence of an event and the report of the respondent about the event, even when long-term coverage involves longitudinal research as, for example, in the investigations by Caspi et al. (2003). It has long been known that simple forgetting is not the only problem (e.g., Blaney, 1986). Rather, the reporting of events on traditional checklists is susceptible to recall biases (e.g., Cohen, Towbes, & Flocco, 1988; Raphael & Cloitre, 1994).

A methodological study by Southwick et al. (1997) suggested that even the most objective-seeming checklist descriptions of combat events are susceptible to such bias. These

investigators examined responses by 59 United States Gulf War veterans of combat experiences to a 19-item "Desert Storm Trauma Questionnaire," which was administered both at 1 month after the veteran's return to the United States (Time 1) and 2 years later (Time 2). At Time 2 testing, the large majority of the veterans, 88%, changed their responses to at least one of the 19 checklist items; fully 61% changed two or more responses. Most important, those with more symptoms of posttraumatic stress disorder (PTSD) at Time 2 were more likely to report an increase in number of combat experiences from Time 1 to Time 2. A similar study with similar results was conducted by Roemer, Litz, Orsillo, Ehlich, and Friedman (1998) of 460 U.S. soldiers who had served in the peacekeeping mission in Somalia. Both sets of investigators note that the apparent dose/response relationship between exposure and PTSD symptoms was attributable in part to recall bias.

Two more recent studies focused on PTSD in war veterans have found less recall bias: one with Gulf War veterans (King et al., 2000), the other with Vietnam War veterans (Koenen, Stellman, Stellman, & Dohrenwend, 2000). A third study, with members of a Netherlands peacekeeping force in Cambodia, found almost none (Bramsen, van der Ploeg, Dirkzwager, & van Esch, 2001). How serious, then, is the problem of recall bias?

It is likely that the size of the effect varies directly with the rate of clinically significant PTSD in the sample (very low in the study of Netherlands peacekeepers by Bramsen et al., 2001), the volatility of PTSD between the first and second recall measures, and changes in the climate of public opinion about the war or mission in which the exposure took place (Koenen et al., 2000; Morgan, 1997). The fact remains, however, that in none of these studies is there an independent, objective measure of what actually happened at the time of the exposure against which to test respondent recall at the time of first questioning. Therefore it is impossible to tell in absolute terms how large the bias is in subsequent recall at the Time 1 questioning.

With the size of the initial systematic bias between actual occurrence and respondent recall at Time 1 unknown, these results showing systematic recall bias between the Time 1 follow-up and the Time 2 follow-up reports by the respondent are cause for concern, whether this bias appears relatively large or relatively small as measured by changes of recall per se. The need for attention to the problem of recall bias is heightened when, as is the case here, the bias can produce misleading results on an issue as important as the presence and strength of a dose/response relationship between exposure and PTSD. That systematic recall bias of this kind occurs at all between successive measures of postexposure recall indicates that state-dependent or mood-congruent recall processes (e.g., Kihlstrom, Eich, Sandbrand, & Tobias, 2000) are operating. Reliance on measures of exposure susceptible to such processes compromises tests of the dose/response relationship between exposure and PTSD.

These recall bias studies provide clues as to where this susceptibility may lie, and these clues lead back to the problem of intracategory variability. The categories on the checklists used in the research on recall bias were thought to be objective and to measure highly traumatic events. Table 3 contains examples of items that showed substantial amounts of change between Time 1 and Time 2 in the study by Southwick et al. (1997).

Morgan (1997) reinterviewed 36 of the Gulf War veterans in this study and confronted them with the discrepancies. He found that the veterans were surprised and could supply no coherent explanation of why they changed their responses. Morgan concluded that the veterans were reinterpreting the meaning of terms in the checklist items in the context of contemporary situations, including respondents having obtained, from the media and from discussions with other veterans, additional and often different information about the war during the 2-year interval. This reinterpretation may be because different sources of information are discounted in memory while the information itself is retained, resulting in reconstructions of events that

bear little resemblance to actual experience at the time of the events (Tourangeau, 2000). If so, then the more symptomatic individuals are the ones who are most likely to select information that increases their reports of negative events.

The results from investigations of recall bias suggest that traditional checklists of events have problems not only of intracategory variability between individual respondents, as was shown earlier, but also problems of intracategory variability for the same individual over time. The breadth of their categories, which leave traditional checklists open to differing interpretations by respondents of the types of experiences a category can include, may make traditional inventories particularly susceptible to systematic recall biases. If this is the case, then procedures for reducing intracategory variability should also reduce recall biases.

Problems With Criterion Validity—It is sometimes possible to develop indicators of the occurrence of events and their characteristics that are independent of self reports and that can, at least partially, verify the accuracy of the reports. For example, the availability of military and historical records can be used to indicate the probable severity of exposure of U.S. veterans to war zone stressors in Vietnam (Dohrenwend et al., 2004); court records can document childhood victimization (Widom, 1998); and medical records can substantiate events such as abortion (Schaeffer, 2000). For most events on checklist inventories, however, no independent, self-evident, and readily available gold standards exist to establish, independently of self-report, whether a reported event actually occurred, much less its important characteristics. Perhaps the nearest so far to a check on the accuracy of checklist reports in general are studies comparing respondent reports to reports about the respondent by significant others, such as spouses, other relatives, or friends.

Coinformant reports are not, however, ideal checks on the accuracy of respondent reports. An informant may lack first-hand information about some of the events the respondent experienced because the coinformant was not present when the events occurred. The coinformant may either have no hint that such events occurred or may be relying on the respondent's account of the events rather than on the coinformant's own independent observations. However, choice of appropriate coinformants can reduce these problems; for example, peers for some types of events experienced by adolescents and parents for others, depending on the sphere of activity in which the event occurs (Williams & Uchiyama, 1989). Nevertheless, if the respondent and the coinformant are thinking of different actual events when they see a particular broad event category, then they are not going to agree on whether it happened. For example, if the respondent uses the euphemism *laid off* for a job loss and the informant recalls that the respondent was fired for cause, then only the respondent will check the item *laid off* on the checklist; reports of serious physical illness would also differ if a respondent's health problem was viewed as serious by one but not the other.

In view of these possibilities for different interpretations, it is not surprising that reports of coinformants, in samples of psychiatric patients (for reviews, see Neugebauer, 1983; Tennant et al., 1981) and in a sample of nonpatients and patients (Yager, Grant, Sweetwood, & Gerst, 1981), show poor correspondence with respondent reports on traditional checklists. For example, in a study of 18 outpatients with diagnoses of schizophrenia, Neugebauer (1983) reported a mean intrapair agreement for all events on the PERI checklist (Dohrenwend et al., 1978) of .22, with a range from 0 to .42. For their sample of 102 nonpatient men (Department of Veterans Affairs and university employees), Yager et al. (1981) reported agreement for only a third of the events that were reported by either the respondent or significant other on the Holmes and Rahe SRE.

Problems With Construct Validity—Investigators who study environmental influences on health disagree on how to define *stress* (e.g., review by Cohen, Kessler, & Gordon,

1997b). Nevertheless, as Cohen and colleagues point out, all investigators “share an interest in a process in which environmental demands tax or exceed the capacity of the organism, resulting in psychological and biological changes that may put the person at risk . . . [for adverse health outcomes]” (p. 3). Within this general framework of agreement, life events are important representations of environmental demands.

It is generally recognized that the events occur in the situations in which people live their everyday lives (e.g., Dohrenwend, 2000). These situations can vary from the regular activities involved in domestic relationships, education, work, and play that are, by and large, satisfying for most people in most communities in times of peace, on the one hand, to extreme situations of persistent threat, on the other hand. The latter include combat during wartime, prolonged human-made or natural disasters, longstanding domestic violence and child abuse, severe chronic physical illness, and poverty. Note that various demographic factors can be associated with extreme situations but do not, in and of themselves, constitute such situations. For example, low socioeconomic status (SES), which is defined by low education, occupation, and income, is associated with poverty; however, low SES is not synonymous with poverty, which is defined by having inadequate resources to meet the demand of daily living (e.g., O’Hare, 1989). By contrast with the indicators of poverty, the indicators of low SES do not, in and of themselves, define an unusually threatening ongoing situation.

The *extreme situation end* of the continuum of usual activities is characterized far more by chronic stressors or ongoing difficulties that make severe and continuing demands on individuals. Events in such extreme situations need to be assessed not only for their impact on usual activities of the extreme situations but also on the activities in the more usual situations to which the individuals may return—as, for example, the return of a veteran to civilian life after service in a war zone.

In addition, inventorying approaches share the assumption that diverse types of events have characteristics in common that determine their impact. Investigators differ, however, in the characteristics of events they emphasize as making them more or less stressful, severe, or demanding. As noted earlier, for Meyer (1951), as translated by Holmes and Rahe (1967), events are stressful to the extent that they bring about changes in the life of the individual that require his or her readjustment. In accord with this focus on change per se, Holmes and Rahe include positive and negative incidents in their list of stressful events. However, it has become evident that negative events (but not positive events) are positively associated with psychological distress and disorder (e.g., Grant, Sweetwood, Yager, & Gerst, 1981; Lewinsohn, Rohde, & Gau, 2003; Paykel, 1974; Vinokur & Selzer, 1975; Zautra & Reich, 1983). Therefore it is not surprising that investigators, focusing on adverse outcomes involving psychopathology especially, have developed a variety of alternative definitions of the severity of stressful life events that tend to emphasize their negative characteristics. Examples of these characteristics are undesirability (Sarason et al., 1978), objective negative impact (Paykel, 1997), the extent to which the events involve loss of resources (Hobfoll, 1989), their degree of contextual threat (Brown & Harris, 1978), and the extent to which the events are likely to contribute to uncontrollable negative changes in the usual activities of most individuals who experience them (Dohrenwend, 1998, 2000).

Dohrenwend (2000) has posited that at least six general characteristics of events are likely to contribute to the nature and extent of their contribution to uncontrollable negative changes in the usual activities of most individuals who experience them. In addition to *valence* (positive or negative, desirable or undesirable, involving gain or loss), these are *source* (occurrence caused by factors in the external environment in contrast with occurrence resulting from actions of the individual); *unpredictability* (the extent to which the occurrence of event would not be foreseen by most people who experience it); *centrality* (from life threat, to threat to physical

integrity, to threat to basic needs, to threat to other goals); *magnitude* (in terms of likely change in the usual activities of most people who experience the event); and *potential to exhaust the individual physically*.

Unlike the other five characteristics, source is more likely to be associated with differences in the processes that contribute to uncontrollable negative changes than with the magnitude of such changes. When the source of the event is mainly external (e.g., as in death of a loved one), factors in the environment are more likely to be implicated, and the event can be described as *fateful* (Dohrenwend, 1979); when the actions of the individual contribute strongly to the occurrence of the event (e.g., being fired for cause), prior disorder and other personal predispositions, including genetic liability to the occurrence of nonfateful negative events (Kendler, 1998), are more likely to be implicated. Other things equal, no self-evident way comes to mind in which fateful, compared with nonfateful, negative events would contribute more to uncontrollable negative changes in the life of the individual. Much needs to be learned about the conditions under which one or the other type of source is more important in determining the impact of events in relation to various disorder outcomes.

Regardless of the problems in dealing with the complexities of the characteristics of events and the situations in which they occur that will be considered later, a more immediate challenge exists to the construct validity of traditional life event inventories as indicators of *environmental* demands. The challenge resides in the types of broad event categories the inventories include. As Hudgens (1974) pointed out many years ago, the list of 43 event categories covered in the Schedule of Recent Experience (SRE) contains a substantial number of items that, rather than being fundamentally important environmental incidents, could include manifestations of psychiatric disorder; for example, the categories of *major change in eating habits*, *major change in sleeping habits*, *marital separation from mate*, and *being fired from work*. Depending on the circumstances in which they occur, the first two items could be outright symptoms of psychiatric disorder; the second two could be indicators of problems in social functioning related to the presence of psychiatric disorder.

Therefore it is not surprising that such events correlate highly with psychopathology (e.g., Brett, Brief, Burke, George, & Webster, 1990), although they do not appear to account for the correlation between the total checklist events and such outcomes (Zimmerman, O'Hara, & Corenthal, 1984). To the extent that a checklist contains categories that can include symptoms of psychological disorder or disabilities related to such disorder, the measurement of stressful events as a putative antecedent risk factor will be confounded with the measurement of the putative psychiatric outcome—and the potential for such confounding appears considerable in the SRE (Dohrenwend, Dohrenwend, Dodson, & Shrout, 1984). Without more information about the actual events in these categories, it is impossible to tell whether they are fundamentally important environmental incidents or symptoms of psychiatric disorder.

Summary of Impact

The consequences of intracategory variability in traditional checklists include unreliability of recall and, even more serious, susceptibility to recall bias. The latter problem, previously illustrated by the studies of the relation of checklists of exposure to stressful war zone events and PTSD, is especially serious because the susceptibility of checklists of events to recall bias can inflate the supposed dose/response relationship between exposure and a psychopathological outcome. An additional source of confounding of the measurement of exposure occurs when a category such as marital separation or divorce includes events that may be brought about by disorder-related functioning of the respondent, as well as by events with occurrences that are unaffected by the respondent's behavior. This type of intracategory variability undermines the construct validity of the traditional checklist as a measure of environmentally induced stress. Problems of criterion validity exist as well. It is unlikely, for

example, that coinformants will confirm that a respondent has experienced an event on a checklist if the respondent and the informant are thinking of different types of events within the same broad checklist category.

Approaches to Reducing Intracategory Variability in Checklists

Three procedures have been developed to reduce the intracategory variability in traditional checklist approaches. One involves subjective appraisals of stressfulness by the respondent. A second supplies definitions to the respondent, to the interviewer, or to both regarding what is and what is not to be included in an event category. The third involves intensive interviews with the respondent to elicit details of the actual event reported under the various event topics and ratings by the investigators of the characteristics the researchers believe are important.

This third approach is far more labor intensive. Although its purpose, similar to the previous two, is to improve procedures for inventorying stressful life events, the change from relatively economical modifications of traditional checklists on the SRE model to a time-consuming and labor-intensive procedure is of sufficient magnitude to suggest that it be described as something other than a checklist approach. It is referred to here as a *narrative-rating methodology*.

Respondent Appraisals of Stressfulness

Obviously, intracategory variability has resulted in the inclusion of objectively small events and objectively large events in the same-event categories. Some investigators have attempted to reduce this variability economically, by having the respondents provide ex post facto subjective judgments of the severity of negative impact of the events they experienced (e.g., Grant, Gerst, & Yager, 1976; Sarason et al., 1978). This scoring procedure leads to a stronger relationship between events reported and adverse psychological outcomes (e.g., Tennant & Andrews, 1978; Zuckerman, Oliver, Hollingsworth, & Austrin, 1986).

It has long been known, however, that such judgments are influenced by the nature of the outcome in retrospective case/control studies or by antecedent psychological risk factors, such as prior disorder or coping history, in prospective research (cf. Brown & Harris, 1978; Grant et al., 1976; Lennon, Dohrenwend, Zautra, & Marbach, 1990; Raphael & Cloitre, 1994; Schless, Schwartz, Goetz, & Mendels, 1974; Theorell, 1974). For example, Theorell investigated two sets of ratings—by samples of patients with neurotic diagnoses, patients who had myocardial infarctions (MI) and controls—of items on a revised version of the SRE. One rating was of the amount of adjustment the event would require; the other was of the amount of “upsettingness” that the event would cause. Patients with neurotic diagnoses rated both more upsettingness and more adjustment than the controls; the patients with MIs rated more upsettingness than the controls. Information about the events actually experienced in the 12 months before the MI had been obtained from both the patients with MIs and the controls, for a corresponding 12-month period. The ratings of patients who actually experienced the events did not differ from the ratings from patients who did not experience the events; by contrast, controls who experienced the events tended to rate them as less upsetting and requiring less adjustment than the controls who did not experience the events. These results show that the tendency of the patients to appraise the same checklist events as more upsetting than did the controls occurred whether or not the events had been experienced.

It could be argued that use of appraisals that are antecedent to psychiatric outcomes and independent of prior disorder and other personal predispositions might be a viable and economical way to reduce intracategory variability. These are not easy conditions to satisfy. Moreover, even under these conditions of temporal priority and independence from personal predispositions, accurate objective measurement of the events is needed to assess the extent to which the appraisals are commensurate with the objective characteristics of the events or,

rather, vary with other factors such as gender, ethnic/racial background, and SES. Appraisal processes are important in life stress processes (Lazarus & Folkman, 1984). However, it makes more sense to consider appraisals as one of a variety of other relevant variables in life stress processes rather than as economical means of reducing intracategory variability in traditional checklists (e.g., Dohrenwend, 1998, 2000). By doing so, it becomes possible to investigate how appraisals are related to objectively measured events and other important variables in life stress processes.

Specification of Inclusion and Exclusion Criteria for Event Categories

A few investigators have departed from traditional checklists by defining for the interviewer, the respondent, or both what events are to be included in an event category. For example, interviewers using Paykel's (1983) checklist of 61 events are given detailed instructions on administration, including definitions of what is and what is not to be included in each category; they are instructed to probe for enough details to be able to decide whether or not the event should be included in the category. Other examples are life events interviews developed by Wittchen, Essau, Hecht, Teder, and Pfister (1989) and screening questionnaires for potentially traumatic events developed by Goodman, Corcoran, Turner, Yuan, and Green (1998) and Kubany et al. (2000). Zimmerman, Pfohl, and Stangl (1986) have referred to such attempts to increase item specificity as leading to a "second generation" of checklist inventories.

As with traditional checklists, these instruments have good test-retest reliability for total scores over brief periods of time. In addition, however, they tend to have far better test-retest agreement for individual events (e.g., Goodman et al., 1998; Wittchen et al., 1989). An indication also exists of much less fall off of event reporting over time. For example, Wittchen et al. report an average falloff rate of only .36% per month in recall of events measured by a detailed interview and use of memory aids over an 8-year period of recall, which compares favorably with rates such as the 4 to 5% per month over 8 to 9 months for traditional checklists reported by Jenkins, Hurst, and Rose (1979) and Monroe (1982). As would be expected, falloff is least for major negative events (e.g., Raphael et al., 1991; Wittchen et al., 1989).

The checklists developed by Goodman et al. (1998) and by Kubany et al. (2000) to screen for the occurrence over the life course of traumatic events illustrate the kinds of procedures that can be used to specify detail. For example, in their measure to screen for potentially traumatic events, the Stressful Life Events Screening Questionnaire (SLESQ), Goodman et al. (1998) follow positive responses to the broad question, "Has an immediate family member, romantic partner, or very close friend died as a result of accident, homicide, or suicide?" with probes for the respondent's age at the time, how the person died, the respondent's relation to the person lost, and how often the respondent saw or otherwise had contact with the deceased in the year before the death.

The procedure developed by Kubany et al. (2000) is called the Traumatic Life Events Questionnaire (TLEQ). It reduces intracategory variability in magnitude by providing more detailed definitions than most second-generation measures of what is to be included in its categories. It differs from the procedure used by Goodman et al. (1998) in that, instead of using follow-up of positive responses to rather broad questions with probes consisting mainly of the closed questions, the Kubany et al. procedure includes the details of the events in the questions themselves. Examples of these TLEQ items are, "Were you involved in a motor vehicle accident for which you received medical attention or that badly injured or killed someone?" and "While you were growing up, were you physically punished in a way that resulted in bruises, cuts, or broken bones?" These details clearly reduce intracategory variability in the checklist items, and Kubany et al. report much better test-retest reliability, not only for total scores but also for individual events than has been found with traditional checklists.

What is less clear, however, is whether the events included in response to further specification by the types of procedures illustrated by either the SLESQ or the TLEQ will be exhaustive of the major events in each event category. The danger in providing explicit inclusion and exclusion criteria is that such detail may make the definition of events to be reported too narrow. Major events that do not clearly meet such criteria or are construed by the respondent not to meet such criteria may be missed altogether (Goodman et al., 1998). Some evidence suggests that the greater the detailed definition provided for the events in each checklist category, the greater will be the test-retest reliability of the instrument for individual, as well as total, events over brief intervals, the less the falloff in event reporting with length of the recall period, and the greater the agreement between respondent and coinformant (e.g., Paykel, 1987). A price for these improvements, however, may be failure to elicit major events that do not fit neatly into the detailed definitions of what should be included in the checklist category.

Narrative-Rating Methodology

The dominant alternative approach in life event measurement, unlike the checklist approach, is labor intensive and involves the collection and analysis of detailed information about the events reported. It is this narrative information about the event that makes it possible to reduce intracategory variability. For example, if one knows the details of what actually happened, then he or she can distinguish between major events (e.g., those with magnitude ratings of more than a moderate amount of change as illustrated in Table 1) and minor events within a particular category and obtain evidence of the sources of an event that, if identified only by a positive response to a checklist category, may otherwise be ambiguous (as with the category *divorce* that may elicit events precipitated to a greater or lesser extent by the individual) or misleading (as when *laid off* is a euphemism for being fired). Ratings by trained judges can be made of the event characteristics—such as valence, source, and magnitude—that are of interest.

As noted earlier, the best-known example of a narrative-rating method is the Life Events and Difficulties Schedule (LEDS) developed by Brown and colleagues (e.g., Brown & Harris, 1978). LEDS has been used more frequently abroad than in the United States (e.g., Brown & Harris, 1989), but it, or variations of it, have been adopted by a number of U.S. investigators in recent years (Duggal et al., 2000; Frank, Anderson, Reynolds, Ritenour, & Kupfer, 1994; Garber, Keiley, & Martin, 2002; Goodyer & Altman, 1991a, 1991b; Hammen, 1991; Kendler, Hettema, Butera, Gardner, & Prescott, 2003; McQuaid, Monroe, Roberts, Kupfer, & Frank, 2000; Monroe, Kupfer, & Frank, 1992; Rudolph & Hammen, 1999).

Starting with its introduction over 25 years ago (Brown & Harris, 1978), the focus of LEDS has usually been on relatively recent events, that is, those occurring within the past year or less. LEDS requires intensive interviewing about the details of each event that occurred. This instrument was designed to deal with both the problem of intracategory variability in objective scoring of checklist categories and the problem of confounding involved in subjective scoring of the severity or stressfulness of the event by the respondent. Trained raters evaluate the likely “contextual threat” of the event for the individual by assessing its place within the respondent’s “biographically determined circumstances” (Brown & Harris, 1978, p. 90).

Another approach to eliciting detailed information about recent major stressful events is called the Structured Events Probe and Narrative Rating Method (SEPRATE; Dohrenwend et al., 1993). Preliminary versions of this instrument have been used in several studies (Lennon et al., 1990; Mazure, Bruce, Maciejewski, & Jacobs, 2000; Shrout et al., 1989; Stueve, Dohrenwend, & Skodol, 1998). As with LEDS, detailed descriptions or narratives of the event are elicited and rated, not by the respondent, but by trained raters. What is meant by a narrative in this approach is not dissimilar to eliciting a record of the facts through testimony in a court of law, with the aim of developing a “readily digestible account of what happened, how it happened and, if possible, why it happened” (Weitz, 1987, p. 202).

The theoretical framework served by this approach, however, requires that life events measures be distinct from measures of other components of life stress processes (e.g., SES, personal predispositions, and social network supports) so that their interrelations can be investigated and their separate and joint contributions assessed (Dohrenwend, 1998). This requirement leads to a difference with the Brown and Harris LEDES procedure. LEDES combines situational and personal variables that may be important risk factors into the single life event measure of contextual threat. As a result, LEDES provides no way to tell which of the components of life stress processes encompassed in the global threat rating account for a particular association (Kessler, 1997; Tennant et al., 1981; Wethington, Brown, & Kessler, 1997). In the SEPRATE approach, what gets left in a narrative and what is taken out before the ratings depends on the theory underlying the hypotheses to be investigated in a particular study; for example, when the roles of ethnic/racial background and SES are of theoretical interest, information about these variables is removed from event narratives so that their independent contribution can be assessed (Stueve et al., 1998). With SEPRATE, more generally, the emphasis in the narrative material given to the raters is more on the here and now of what happened than on the biographically determined circumstances of the individual.

Narrative-rating procedures have proved far superior to traditional checklist approaches in test-retest reliability for individual events (cf. Gorman, 1993; Paykel, 1983). Moreover, in contrast to the poor agreement with coinformants for traditional checklists noted previously, intrapair agreement with LEDES has been found to range between 78% and 91% (the latter for LEDES events rated as carrying severe contextual threat) in samples with diagnoses of depression and schizophrenia (Brown et al., 1973). A particularly valuable example is the use of narratives from adult sisters to measure childhood abuse and neglect (Bifulco, Brown, Lillie, & Jarvis, 1997). In this study, sisters of similar age provided narratives of their own experiences and the experiences of their sisters. The investigators were able to show that, although there was marked difference in the amount of abuse and neglect experienced by each member of the sister pairs (concordance of about 0.4 as measured by kappa), there was good agreement between the sisters about the abuse and neglect that each experienced (kappa of about 0.7). Such agreement between the sisters increases confidence in the accuracy of the adult sisters' retrospective reports about their childhood that could otherwise be highly suspect.

With regard to construct validity, evidence suggests that narrative-rating procedures provide more interpretable and sometimes (but not always as will be seen later) stronger associations with adverse psychological outcomes than traditional checklist measures (Bebbington, Tennant, Stuart & Hurry, 1984; Duggal et al., 2000; McQuaid et al., 2000; Shrout et al., 1989). When effect sizes are measured by odds ratios or attributable risks, moreover, it is evident that these effect sizes are substantial (Cooke & Hole, 1983). Consider by way of examples some results from two case/control studies that used narrative-rating procedures to measure stressful events in the 3 months before the onset of episodes of major depression. The first study by Ormel, Oldehinkel, and Brilman (2001), with LEDES as the measure of stressful life events, found an odds ratio of 25.9 for one or more severe events in a comparison of case and control groups of older adults from the general population. The second study by Stueve et al. (1998) used SEPRATE to measure stressful events and was particularly concerned with major negative events that were fateful, that is, negative events that occurred independently of the behavior or actions of the individual. The odds ratio was 13.8 for one or more such recent events in a comparison of a sample of psychiatric outpatients with recent episodes compared with controls sampled from the community.

Because the customary procedure with checklists is to sum events rather than dichotomize the measures as in the previously discussed case and control studies using narrative-rating procedures, making a direct comparison of the effect sizes obtained with the two procedures is usually impossible. The second of the two case/control studies, however, does provide an

opportunity for this comparison. This study was conducted in an Upper Manhattan section of New York City and involved a comparison of 404 adults sampled from the general population from whom persons with recent (in the preceding year) episodes of major depression had been removed, and 96 outpatients diagnosed as having recent episodes of major depression. As a result of a concerted effort to oversample them, half of the patients were in their first episode of major depression. Focus is on the following 12 checklist events that had been judged to be negative and likely to be fateful in previous research with the PERI checklist (Dohrenwend, Shrout, Link, Martin, & Skodol, 1986): spouse died; child died; close friend died; miscarriage or stillbirth; found out cannot have children; family member other than spouse or child died; lost home through fire, flood, or other disaster; physically assaulted or attacked; cut in wage or salary without demotion; did not get expected wage or salary increase; laid off; and unable to get treatment for illness or injury. The controls were asked to report events on the checklist that occurred in the year before the interview; the cases were asked to report on events in the year before the onset of the episode of depression.

After filling out the PERI checklist, the respondents were questioned about details of the events. The research team rated the resulting narratives for actual fatefulness and hypothetical amount of negative change. The anchors for the ratings of magnitude are those shown earlier in Table 1. The ratings of source, from which fatefulness was estimated, were made in two parts: the prelude leading up to the event and the immediate occurrence of the event. For example, in an event involving loss of a job, the prelude might consist of repeatedly being late for work and not getting one's job done; the immediate occurrence might be being told by your boss that you are fired. The ratings for this event would be as shown in Table 4.

To be considered fateful, both the prelude to the event and its actual occurrence had to be rated as mostly or completely determined by external circumstances. In making these magnitude and fatefulness ratings, the raters were blind as to whether the event narratives being rated came from patients or nonpatient controls to avoid possible rater bias in favor of the stress hypothesis. Data on ethnic/racial background and educational level were also removed, because these were variables of interest in the research that needed to be assessed for their independent contribution.

Table 5 compares three measures of the 12 negative events as they differentiate cases from controls: one or more unmodified "fateful" checklist events; one or more events verified to be fateful; and one or more events verified to be fateful and likely to entail more than a little change (Shrout et al., 1989). As Table 5 shows, the odds ratios are much larger for the measure revised for both fatefulness and magnitude on the basis of detailed interview data—even after scores on a scale of nonspecific distress or demoralization were entered to control for mood-congruent recall bias.

By contrast with traditional checklists, labor-intensive procedures have also led to important findings suggesting, as noted earlier, that among inventories of recent events, it is major negative events of substantial magnitude that are most strongly related to important types of psychiatric disorders, especially first onsets of major depression, and that these relationships are substantial (e.g., Brown & Birley, 1968; Brown & Harris, 1978; Ormel et al., 2001; Stueve et al., 1998). These findings on the importance of major negative events are consistent with the results of case studies of individual events, such as spousal bereavement (Clayton, 1998), divorce (Bruce, 1998), rape (Kilpatrick, Resnick, Saunders, & Best, 1998), child abuse and neglect (Widom, 1998), and serious physical illness (Dew, 1998). Moreover, it has been possible to show that these major events are much more likely to be fateful for depression than for schizophrenia (Stueve et al., 1998).

No studies were found in which narrative-rating procedures and second-generation checklists were administered to the same respondents. However, in at least two studies in which both LEDS and traditional checklist measures have been administered to the same respondents, the two methods have shown equal ability to discriminate depressed patients from nondepressed controls. (Duggal et al., 2000; Faravelli & Ambonetti, 1983). This similarity occurs despite the fact that, when traditional checklist and narrative-rating instruments are administered to the same respondents, they tend to identify different events (Costello & Devins, 1988; Duggal et al., 2000; Katschnig, 1986; Raphael et al., 1991).

A reason for the parity in predictive power in the two methods in these studies may lie in the relation of minor negative events to a history of recurrences of depression. Most episodes of major depression in patient samples and cross-sectional community samples tend to be recurrences (e.g., Kessler, 1997). The inpatient cases of severe major depression in the Faravelli and Ambonetti (1983) study probably had a high ratio of frequently recurrent rather than first-episode major depression at the time of the research. The episodes of major depression in a substantial minority of the adolescent patients in the study by Duggal et al. (2000) were recurrences. As previously noted, minor events have been shown in several studies to be positively associated with recurrences of major depression; by contrast, first episodes of depression are more likely to require the presence of major negative events (Hammen, 2005; Monroe & Harkness, 2005). It is possible that the equal predictive power of the traditional checklists found in some studies is related to the likelihood that such checklists contain a higher proportion of relatively minor incidents, many of which, according to the results in Table 2, would not meet even minimal inclusion criteria for what constitutes a stressful life event in studies using narrative-rating procedures; for example, be likely to produce negative changes in usual activities lasting more than 1 week for most people who experience the event (Dohrenwend et al., 1990; 1993) or to involve contextual threat lasting more than 1 week (Brown & Harris, 1978).

Monroe and Harkness (2005) point out that the combination of the decreasing importance of major stressful events and the increasing importance of minor events in recurrences of major depression is consistent with hypotheses derived from “kindling” (Post, 1992) and other stress-sensitization theories in contrast to “autonomy” theories in which stressful events are held to become unimportant for recurrent episodes. Further investigation of these intriguing hypotheses requires differentiating between major and minor negative events. To do so, it will be necessary to use methods that accurately assess the magnitude of negative events ranging from those that are catastrophic to small incidents or hassles (Kanner et al., 1981; Zautra et al., 1986). Given the problems of recall for minor events discussed previously and the tendency of daily hassles to be confounded with symptom outcomes (Dohrenwend & ShROUT, 1985; Dohrenwend et al., 1984) and to require diary methods that are rarely used for more than a few weeks (Eckenrode & Bolger, 1997), tests of the role of the full range of major and minor events in first onsets and recurrent depression will have to involve innovative integration of what have been, in the past, separate lines of research. Clearly, the problem of intracategory variability compromises the construct validity of traditional checklists for such purposes.

Summary of Procedures for Reducing Intracategory Variability

One of the three procedures for reducing intracategory variability can cause more problems than it solves. This is the use of retrospective subjective appraisals of stressfulness, upsettingness, or some similar subjective perception to assess the magnitude or severity of the event. Such perceptions have been highly likely to be affected by the psychiatric outcome. They are perhaps the surest way to confound the measure of exposure and the measure of the health outcome. Two far more defensible procedures exist. One is to reduce intracategory variability by spelling out inclusion and exclusion criteria for the events to be included in the

category. More or less intensive attempts to do this have made their way into a second generation of checklists that have generally been found to be more reliable than traditional checklists. The problem with this procedure is that the more detailed the specification of the inclusion and exclusion criteria, the more likely it is that events that are important will be missed because they do not meet these particular criteria. The second more effective procedure involves labor-intensive interviewing and narrative ratings. This procedure is designed to elicit accounts of the events experienced in sufficient detail so that trained investigators can rate the important characteristics of the events. Narrative-rating instruments provide large gains in reliability and validity in the measurement of major stressful events. Unlike checklist measures, however, this procedure is time-consuming and expensive. The first, best known, and most widely used of these instruments, LEDES, was developed over 25 years ago. It has been estimated that it takes an average of 16 hours to conduct and rate one LEDES interview about recent events (Wethington et al., 1997), and most studies have many variables to measure in addition to life events.

Problem of Economy

The problem of economy with intensive interview and narrative-rating procedures should not be underestimated. Despite the manifest superiority in reliability and validity of narrative-rating procedures, these labor-intensive procedures are rarely used. Grant et al. (2004) point out, for example, that of the 500 studies that met inclusion criteria for their reviews of more rigorous research on life events and psychopathology in children and adolescents since 1987, less than 2% used labor-intensive interview procedures to measure stressors. As in research with adults, the overwhelming majority of these studies relied on economical self-report checklists that, although constructed by many different investigators, are generally of the traditional type with broad event categories (Grant et al., 2004). The result is that there do not now exist methods of choice that are likely to be used to take important next steps in research on stressful life events as risk factors for various types of psychopathology over the life course.

As examples of important steps, Grant et al., (2004) point to the need to collect reliable and valid normative data on the occurrence of life events and to construct theoretically or empirically based taxonomies of the most important types and characteristics of events as cornerstones of a foundation for further progress (Grant et al., 2004). A particularly valuable taxonomy might start with two main axes. One axis would consist of types of situations ranging from usual to extreme; the other axis would consist of types of groups demarcated by age, gender, ethnic/racial background, and SES. This formulation would have the virtue of using demographic distinctions that are strongly associated with differences in rates of important types of psychopathology (e.g., Kohn, Dohrenwend, & Mirotnik, 1998), as well as with differences in types of life events at each hierarchical level of centrality—life threat, threat to physical integrity, threat to basic necessities, and threat to important goals (Dohrenwend, 2000). The availability of such a classification system and normative data on the distribution of important events within it would make it possible to identify and follow over time groups at high risk of exposure to severe environmental stressors. If the past is any guide to the future, however, advances of this kind are unlikely to be made unless economical, as well as reliable and valid, procedures for inventorying stressful life events over the life course become available.

Toward an Economical Resolution of the Problem of Intracategory Variability

Economy is, of course, a matter of degree. The most economical procedure is the self-administered questionnaire (SAQ). SAQs do away with the expenses of training interviewers and paying them to conduct interviews about life events and may not lose anything in reliability when compared with interview-administered versions of the same questionnaire (e.g., Kubany et al., 2000). Moreover, some evidence exists that SAQs also have the particular virtue of

eliciting more reports of sensitive events (e.g., events involving matters such as child abuse, rape, and abortion) than personal interviews with the same checklist (Schaeffer, 2000). The results are not, however, entirely consistent, and Schaeffer has speculated that the underlying benefit of more complete reporting derives from privacy that may be ensured in other ways—for example, by computer-assisted interviewing that does not require the presence of an interviewer or a certain level of reading skills on the part of the respondent (Turner et al., 1998), by emphasis on the privacy of the data collection situation and the importance and legitimacy of the research (Schaeffer, 2000), or by both.

It is hard to imagine how self-administered questionnaires would ever be useful in the exploratory stages of research on stressful life events. Moreover, for investigators who need to be able to retrospectively date the occurrence of events in relation to the occurrence of episodes of disorder over substantial periods of time, it is difficult to see how procedures to aid autobiographical memory (e.g., Bradburn, 2000; Tourangeau, 2000), such as the use of life calendar methods (e.g., Belli, 1998; Caspi et al., 1996; Lyketsos, Nestadt, Cwi, Heithoff, & Eaton, 1994), can be implemented without the active presence of an interviewer.

Narrative-rating procedures that require the presence of highly trained interviewers and judges are at the other extreme from SAQs on the economy continuum. The benefits in reliability and validity of narrative-rating procedures are now widely recognized along with their lack of economy and, as would be expected under these circumstances, there have been and continue to be attempts to retain at least some of the benefits of narrative-rating methods with more economical procedures.

At least two attempts are underway to approximate contextual threat ratings of narrative material from intensive semistructured interviews with measures based on more structured interview approaches. The latter approaches place greater reliance on relatively economical closed questions and closed follow-up probes compared with the more open-ended LEDS approach. One of these more structured substitutes for LEDS is an interview called the Structured Life Events Inventory (SLI) that makes more extensive use of closed questions and closed probes and that requires less interviewer training than the more open-ended questioning procedures of LEDS (Wethington et al., 1997). The SLI was investigated in a study of 243 community respondents, half of whom were interviewed with the SLI and half with LEDS. Raters of the SLI interviews were said to reliably distinguish between events representing severe contextual threat and more minor events, and to identify recent (previous 3 months) severe events and difficulties that showed similar positive associations with onset of depressive episodes, as did the ratings based on LEDS. Moreover, use of the SLI for these purposes reduced interview and rating time to an average of 9 hours per respondent compared with the 16-hour average for LEDS. However, only 41 respondents were interviewed with both methods, which must have greatly limited the opportunity to examine the extent to which the SLI and LEDS agreed in the measurement of contextual threat and other characteristics of relatively recent stressful events. Moreover the SLI, though more economical than LEDS, still involves a fairly intensive interview and narrative-rating process.

The other approach, now in its initial stage of development, is an attempt to go a step beyond the SLI in economy by substituting fully structured questions and probes and mechanical scoring for intensive semistructured interviews and narrative ratings altogether (Grant et al., 2004). In addition, by contrast with the SLI, the focus is on children and adolescents rather than adults. The procedure involves analyzing contextual threat ratings based on narratives elicited by intensive semistructured interviews. The purpose of the analysis is to identify the specific items of information in the narratives on which the ratings are based. Once the information is identified, the next step would be to develop structured, closed questions, the

answers to which would provide direct indicators of the nature and severity of the threat posed by the different types of events reported.

Unlike the previously discussed attempts to supplant labor-intensive procedures directly with more economical methods of doing the same thing, there have been attempts to increase economy by developing screening procedures designed to reduce the number of events for labor-intensive investigation. In this approach, minor events are assumed to be relatively unimportant and can be economically screened out so that larger events can be investigated with more intensive interview procedures (Brugha & Cragg, 1990; Costello & Devins, 1988; Goodman et al., 1998; Kubany et al., 2000; Miller & Salter, 1984; Wittchen et al., 1989).

The results of most of the screening studies that have actually conducted intensive follow-up interviews suggest that, if the screening instrument includes a few mandatory open-ended or closed-question probes to elicit more information about context, then the occurrence of the large majority of the events that are positive on the screening instrument will later be verified by more intensive interviews with free probing for details of what actually occurred. This seems to hold for both traditional checklist screens composed of broad event categories (Brugha & Cragg, 1990; Miller & Salter, 1984) and for screens composed of both broad and more specifically defined checklist items focused on traumatic events (Goodman et al., 1998).

The most serious limitation of most of these tests of event screening instruments is that only positive responses on the screening instruments have been followed up with intensive interviews. Such studies cannot therefore address the question of how many important events are being missed by the screens. The exception is the study by Goodman et al. (1998), in which a subsample of the respondents that included screen negatives and screen positives were followed 2 weeks later with an intensive interview by a clinician blind to the results of the earlier screen. This 40-minute interview, which drew questions from a variety of previous studies but is not otherwise described, was designed to cover more intensively the same topics as those addressed with the screening instrument. The test-retest correlation for total events was .77; the median kappa for individual events was .64, with kappa for six of the items falling below .60. The interview elicited more events than the screening instrument, especially with regard to the six low kappa events. These results suggest that screening instruments of this type, despite showing substantial agreement with intensive interviews, will err on the side of underinclusiveness.

As Paykel (1983) pointed out many years ago, all detailed interview approaches start with lists of life event categories. This includes LEDS and SEPRATE. The difference between checklists and intensive interview approaches is that the latter probe positive responses to the listed topics for details to develop narratives of what occurred, with the goals of reducing intracategory variability by permitting identification in the narratives of the events and event characteristics of interest. The resulting measures of the events and their characteristics are obtained by ratings of the narratives made by the investigators. How much detail is needed, however, and how best to provide it are matters for investigation. The question of whether trained raters must obtain the resulting measures or whether more economical mechanical scoring can be used must also be answered. Here is how such investigation might proceed.

Checklist measures would be redesigned to reduce the indeterminate mix of major and minor events within each category. Clues in the research reviewed so far suggest that this could be economically done in one of two ways: (1) adding a few closed-question probes after a positive response to a broad checklist category, as was done by Goodman et al. (1998); or (2) building inclusion or exclusion criteria (or both) into the checklist category itself (Grant et al., 2004; Kubany et al., 2000). In either case, scoring of the responses to the closed-question probes could be done mechanically.

Reducing intracategory variability economically by either of these procedures—(a) developing structured follow-up probes of broad event categories or (b) providing detailed definitions with which to narrow event categories—requires having a great deal of information about the nature of each type of event of interest. The procedures being investigated by Wethington et al. (1997) in the SLI and by Grant et al. (2003) involve searches for information in narratives that is strongly associated with contextual threat ratings made according to LEDS types of inventorying procedures in particular studies. Other more generalizable sources can be found in the vast literature reporting investigations of individual events, such as human-made and natural disasters (e.g., Giel, 1998), bereavement (e.g., Clayton, 1998), marital separation and divorce (e.g., Bruce, 1998), rape (e.g., Kilpatrick et al., 1998), and unemployment (e.g., Kasl, Rodriguez, & Lasch, 1998). Kubany et al. (2000) seem to have drawn, for example, on studies of various types of sexual abuse to specify inclusion and exclusion criteria for their screening items for these types of event.

How much can be accomplished with economically closed questions and closed probes to identify stressful events and measure their important characteristics needs to be empirically investigated. This investigation could be done by comparing and contrasting types of checklist approaches on their ability to screen the events of interest and to measure their important characteristics in samples of respondents from relevant populations of interest. If the interest is in major stressful events, then the focus could be on the occurrence of such events over the full life course. For life course coverage, personal interviews would be indicated, because dating, preferably with the aid of life calendar procedures, would be required. A field experiment designed to test the contrasting screening measures could involve data collection in three phases of increasingly more labor-intensive methods as shown in Figure 1.

Two checklist approaches with different strengths and weaknesses would be used as screening instruments in the first phase of the comparison in Figure 1. One screening instrument, Check I, would be a traditional checklist consisting of broad, overinclusive event categories that would be made more specific by the use of closed-question probes of positive responses to determine, for example, whether the events being reported were major or minor (see earlier examples of the probes used by Goodman et al., 1998, for this purpose). The other would be an underinclusive second-generation checklist made up of fully structured questions that narrow the event categories by spelling out, in more detail than most second-generation checklists, inclusion and exclusion criteria in the checklist items themselves (see previous examples from Kubany et al., 2000). To guard against missing major events that do not clearly meet these inclusion and exclusion criteria, this procedure would make use of “other” categories to gain information about additional events of each type. More specifically, the instrument would consist of stem questions that would be expanded by liberal use of other categories for events similar in some but not all ways to the types of events described by the detailed stem questions. This instrument would be called Stem I. The responses to the closed questions in both fully structured probes in Check I and closed questions in Stem I would be mechanically scored for each event and its important characteristics. The contrasting screening instruments would be randomly alternated between two groups of respondents as shown in Figure 1. This would permit an investigation of how each instrument related to the other, and tests of which instrument more closely approximated criterion Phase 3 measures of major negative events and their important characteristics. The Phase 3 measures would be based on full-structured interview and narrative-rating procedures.

About 400 respondents would be needed to ensure sufficient statistical power to detect carry-over effects and possible interactions with such demographic differences as those involved in gender. All respondents would be interviewed with both Check I and Stem I instruments in Phase 1 of this experimental test, using a cross-over design with a 1-week delay between the administering of instruments to reduce order or carry-over effects. In Phase 2 of the experiment,

one designed to investigate test-retest reliability, the respondents would be reinterviewed about 2 weeks later. At the end of the retest, the respondents would be asked two brief open-ended questions about each of her or his positive responses on either Check I or Stem I (i.e., What happened? and What led up to it?). The purpose of these questions would be to investigate further whether this relatively modest addition would lead to a large increase in accuracy by providing additional detail for measuring important characteristics such as the source, centrality, and magnitude of the event. The Phase 3 criterion measures would be applied to all positive responses elicited by Check I and Stem I about 1 month earlier.

The field experiment summarized in Figure 1 is designed for the purposes of (a) comparing the major stressful events and their characteristics identified by each screening instrument to the major stressful events and their characteristics identified by the other screening instrument, as well as (b) testing the ability of each screening instrument to identify the major stressful events and their important characteristics elicited and measured by the full intensive interview and narrative ratings from Phase 3. The results would be relevant for retrospective research over the life course or long periods of the life course. As noted earlier, such research must rely on personal interviews and life calendar aids to recall. Additional tests of the two types of screening instruments could be developed for different purposes; for example, the study of minor and major events longitudinally in relation to the onset and recurrent course of major depression. If particularly sensitive events were to be investigated, especially with children or adolescents (Turner et al., 1998), then it would be well to test the relative accuracy of computer-assisted interviews, self-administered questionnaires, and personal interview versions of Check I and Stem I.

Conclusions

A tremendous increase in research on stressful life events has taken place since the publication of the Holmes and Rahe checklist in 1967. Despite much criticism, especially with regard to their use in research on psychopathology, such economical inventories composed of broad event categories or topics have proliferated and remain dominant. The criticisms center on problems of reliability and validity that exist within these traditional inventories. As the analyses reported here have shown, the problems with these measures can be traced to a major flaw—the intracategory variability of actual events reported in the broad checklist categories.

The main alternative approach, unlike traditional checklists, involves intensive, time-consuming, and expensive semistructured interviews to obtain detailed narrative information about life events that facilitates judgments by trained raters of the important characteristics of the events. These narrative-rating procedures are far superior to traditional checklists in reliability and validity. However, the gains are bought at tremendous loss in economy, and narrative-rating measures are rarely used.

More economical instruments need to be developed that nevertheless effectively deal with the problem of intracategory variability. Clues about how to do so are found in studies that compare traditional checklist with narrative-rating procedures and in a second generation of checklist approaches that set forth inclusion and exclusion criteria for the event categories of interest. Designs such as the one summarized in Figure 1 above could investigate experimentally how closely redesigned traditional and second-generation inventories can approximate narrative-rating measures in terms of reliability and validity. The purpose of such research would be to develop and improve screening measures that could be substituted for the inadequate checklist measures that have been used in most of the research on stressful life events and psychiatric disorders to date. The screening measures could be constructed to vary in the amount of time-consuming narrative detail they obtained to supplement their otherwise fully structured formats. The most economical measures would contain no open-ended probes and could be

scored mechanically. The increasingly less economical measures would be constructed at points on a progression to full intensive interview and narrative-rating procedures.

The results should facilitate the investigation of an increasing number of important theoretical questions and hypotheses about factors in the onset and course of various types of psychopathology that require complex research designs with many variables in addition to stressful life events. Some examples are questions about the differing roles of major and minor negative events in the onset, by contrast with the course of major depression (Hammen, 2005; Monroe & Harkness, 2005) and, perhaps, other episodic disorders; the issue of the primacy of the stressor in PTSD and dose/response relationships that bear importantly on this issue (Dohrenwend, 1998, 2000); the conditions under which, and the outcomes for which, fateful or nonfateful negative events have greater impact; and, most comprehensively, the nature of gene and environment interactions in the development of psychiatric disorders (Caspi et al., 2003). By knowing what would be gained and what would be lost by substituting relatively brief screening procedures for intensive interview and narrative-rating methods of measuring major stressful events, investigators in such research could focus on what was crucially important to their hypotheses and make informed decisions about how much economy they could afford and still maintain the scientific integrity of the enterprise.

References

- Bebbington P, Tennant C, Stuart E, Hurry J. The domain of life events: A comparison of two techniques of description. *Psychological Medicine* 1984;14:219–222. [PubMed: 6709788]
- Belli RF. The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory* 1998;6:383–406. [PubMed: 9829098]
- Bifulco A, Brown GW, Lillie A, Jarvis J. Memories of childhood neglect and abuse: Corroboration in a series of sisters. *Journal of Child Psychology & Psychiatry* 1997;38:365–374. [PubMed: 9232482]
- Blaney PH. Affect and memory: A review. *Psychological Bulletin* 1986;99:229–246. [PubMed: 3515383]
- Bradburn, NM. Temporal representation and event dating. In: Stone, AA., et al., editors. *The science of self-report: Implications for research and practice*. Mahwah, NJ: Erlbaum; 2000. p. 49-61.
- Bramsen I, van der Ploeg H, Dirkzwager A, van Esch S. Consistency of self-reports of traumatic events in a population of Dutch peacekeepers: Reason for optimism? *Journal of Traumatic Stress* 2001;14:733–740. [PubMed: 11776420]
- Breslau N. Epidemiologic studies of trauma, posttraumatic stress disorder, and psychiatric disorders. *The Canadian Journal of Psychiatry* 2002;47:923–929.
- Breslau N, Davis GC, Andreski P, Peterson E. Traumatic events and posttraumatic stress disorder in an urban population of young adults. *Archives of General Psychiatry* 1991;48:216–222. [PubMed: 1996917]
- Breslau N, Davis GC, Peterson EL, Schultz L. Psychiatric sequelae of posttraumatic stress disorder in women. *Archives of General Psychiatry* 1997;54:81–87. [PubMed: 9006404]
- Breslau N, Kessler RC, Chilcoat HD, Schultz LR, Davis GC, Andreski P. Trauma and posttraumatic stress disorder in the community. *Archives of General Psychiatry* 1998;55:626–632. [PubMed: 9672053]
- Brett JF, Brief AP, Burke MJ, George JM, Webster J. Negative affectivity and the reporting of stressful life events. *Health Psychology* 1990;9:57–68. [PubMed: 2323329]
- Brown, GW. Meaning, measurement, and stress of life events. In: Dohrenwend, BS.; Dohrenwend, BP., editors. *Stressful life events: Their nature and effects*. New York: Wiley; 1974. p. 217-243.
- Brown GW, Birley JLT. Crises and life changes and the onset of schizophrenia. *Journal of Health and Social Behavior* 1968;9:203–214. [PubMed: 5676853]
- Brown, GW.; Harris, TO. *Social origins of depression: A study of psychiatric disorder in women*. New York: Free Press; 1978.
- Brown, GW.; Harris, TO., editors. *Life events and illness*. New York: Guilford Press; 1989. p. 1989

- Brown GW, Sklair F, Harris TO, Birley JLT. Life events and psychiatric disorders: I. Some methodological issues. *Psychological Medicine* 1973;3:74–78. [PubMed: 4692493]
- Bruce, ML. Divorce and psychopathology. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 219-234.
- Brugha TS, Cragg D. The list of threatening experiences: The reliability and validity of a brief life events questionnaire. *Acta Psychiatrica Scandinavica* 1990;82:77–81. [PubMed: 2399824]
- Caspi A, Moffitt TE, Thornton A, Freedman D, Amell JW, Harrington H, et al. The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research* 1996;6:101–114.
- Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, et al. Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science* 2003;301:386–389. [PubMed: 12869766]
- Clayton, PJ. Model of stress. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 96-110.
- Cohen L, Towbes L, Flocco R. Effects of induced mood on self-reported life events and perceived and received social support. *Journal of Personality and Social Psychology* 1988;55:669–674. [PubMed: 3193354]
- Cohen, S.; Kessler, RC.; Gordon, LU., editors. *Measuring stress: A guide for health and social scientists*. New York: Oxford University Press; 1997a.
- Cohen, S.; Kessler, RC.; Gordon, LU. Strategies for measuring stress in studies of psychiatric and physical disorders. In: Cohen, S.; Kessler, RC.; Gordon, LU., editors. *Measuring stress: A guide for health and social scientists*. New York: Oxford University Press; 1997b.
- Cooke DJ, Hole DJ. The aetiological importance of stressful life events. *British Journal of Psychiatry* 1983;143:397–400. [PubMed: 6626860]
- Costello CG, Devins GM. Two-staged screening for stressful events and chronic difficulties. *Canadian Journal of Behavioural Science* 1988;20:85–92.
- Creed, F. Stress and psychosomatic disorders. In: Goldberger, L.; Breznitz, S., editors. *Handbook of stress*. 2nd ed. New York: The Free Press; 1993. p. 496-510.
- Davidson JRT, Hughes D, Blazer DG, George LK. Posttraumatic stress disorder in the community: An epidemiological study. *Psychological Medicine* 1991;21:713–721. [PubMed: 1946860]
- Day R. Life events and schizophrenia: The “triggering” hypothesis. *Acta Psychiatrica Scandinavica* 1989;64:97–122. [PubMed: 7032227]
- Dew, MA. Psychiatric disorder in the context of physical illness. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 177-218.
- Dohrenwend, BP. Problems in defining and sampling the relevant population of stressful life events. In: Dohrenwend, BS.; Dohrenwend, BP., editors. *Stressful life events*. New York: Wiley; 1974. p. 275-310.
- Dohrenwend, BP. Stressful life events and psychopathology: Some issues of theory and method. In: Barrett, JF.; Rose, RM., et al., editors. *Stress and mental disorder*. New York: Raven Press; 1979. p. 1-15.
- Dohrenwend, BP. Theoretical integration. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 539-555.
- Dohrenwend BP. The role of adversity and stress in psychopathology: Some evidence and its implications for theory and research. *Journal of Health and Social Behavior* 2000;41:1–19. [PubMed: 10750319]
- Dohrenwend BP, Link BG, Kern R, ShROUT PE, Markowitz J. Measuring life events: The problem of variability within event categories. *Stress Medicine* 1990;6:179–187.
- Dohrenwend BP, Neria Y, Turner JB, Turse N, Marshall R, Lewis-Fernandez R, et al. Positive tertiary appraisals and post-traumatic stress disorder in U. S. male veterans of the war in Vietnam: The roles of positive affirmation, positive reformulation, and defensive denial. *Journal of Consulting and Clinical Psychology* 2004;72:417–433. [PubMed: 15279526]
- Dohrenwend, BP.; Raphael, KG.; Schwartz, S.; Stueve, A.; Skodol, AE. The structured event probe and narrative rating method (SEPRATE) for measuring stressful life events. In: Goldberger, L.; Breznitz, S., editors. *Handbook of stress: Theoretical and clinical aspects*. 2nd ed. New York: Free Press; 1993. p. 174-199.

- Dohrenwend BP, ShROUT PE. "Hassles" in the conceptualization and measurement of life stress variables". *American Psychologist* 1985;40:780–785.
- Dohrenwend, BP.; ShROUT, PE.; Link, BG.; Martin, JL.; Skodol, AE. Overview and initial results from a risk-factor study of depression and schizophrenia. In: Barrett, JE.; Rose, RM., editors. *Mental disorders in the community*. New York: The Guilford Press; 1986.
- Dohrenwend, BS.; Dohrenwend, BP., editors. *Stressful life events: Their nature and effects*. New York: Wiley; 1974.
- Dohrenwend BS, Dohrenwend BP, Dodson M, ShROUT PE. Symptoms, hassles, social supports and life events: The problem of confounded measures. *Journal of Abnormal Psychology* 1984;93:222–230. [PubMed: 6725756]
- Dohrenwend BS, Krasnoff L, Askenasy AR, Dohrenwend BP. Exemplification of method for scaling life events: The PERI life events scale. *Journal of Health and Social Behavior* 1978;19:205–229. [PubMed: 681735]
- Duggal S, Malkoff-Schwartz S, Birmaher B, Anderson BP, Matty MK, Houck PR, et al. Assessment of life stress in adolescents: Self-report versus interview methods. *Journal of the American Academy of Child and Adolescent Psychiatry* 2000;39:445–451. [PubMed: 10761346]
- EckenrodeJ BolgerN 1997 Daily and within-day event measurement Cohen S Measuring stress: A guide for health and social scientists 80101 New York Oxford University Press
- Faravelli C, Ambonetti A. Assessment of life events in depressive disorders: A comparison of three methods. *Social Psychiatry* 1983;18:51–56. [PubMed: 6867815]
- Frank E, Anderson B, Reynolds CF III, Ritenour A, Kupfer DJ. Life events and the research diagnostic criteria endogenous subtype. A confirmation of the distinction using the Bedford College methods. *Archives of General Psychiatry* 1994;51:519–524. [PubMed: 8031224]
- Garber J, Keiley MK, Martin NC. Developmental trajectories of adolescents' depressive symptoms: Predictors of change. *Journal of Consulting and Clinical Psychology* 2002;70:79–95. [PubMed: 11860059]
- Giel, R. Natural and human-made disasters. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 66–76.
- Goodman LA, Corcoran C, Turner K, Yuan N, Green BL. Assessing traumatic event exposure: General issues and preliminary findings for the stressful life events screening questionnaire. *Journal of Traumatic Stress* 1998;11:521–542. [PubMed: 9690191]
- Goodyer IM, Altman PME. Lifetime exit events and recent social and family adversities in anxious and depressed school-age children and adolescents—I. *Journal of Affective Disorders* 1991a;21:219–228. [PubMed: 1829743]
- Goodyer IM, Altman PME. Lifetime exit events and recent social and family adversities in anxious and depressed school-age children and adolescents—II. *Journal of Affective Disorders* 1991b;21:229–238. [PubMed: 1829744]
- Gorman DM. A review of studies comparing checklist and interview methods of data collection in life event research. *Behavioral Medicine* 1993;19:66–73. [PubMed: 8280964]
- Grant I, Gerst M, Yager J. Scaling of life events by psychiatric patients and normals. *Journal of Psychosomatic Research* 1976;20:141–149. [PubMed: 1271314]
- Grant I, Sweetwood HL, Yager J, Gerst M. Quality of life events in relation to psychiatric symptoms. *Archives of General Psychiatry* 1981;38:335–339. [PubMed: 7212965]
- Grant KE, Compas BE, Stuhlmacher AF, Thurm AE, McMahon SD, Halpert JA. Stressors and child and adolescent psychopathology: Moving from markers to mechanisms of risk. *Psychological Bulletin* 2003;129:447–466. [PubMed: 12784938]
- Grant KE, Compas BE, Thurm AE, McMahon SD, Gipson PY. Stressors and child and adolescent psychopathology: Measurement issues and prospective effects. *Journal of Clinical Child and Adolescent Psychology* 2004;33:412–425. [PubMed: 15136206]
- Gunderson, EKE.; Rahe, RH., editors. *Life stress and illness*. Springfield, IL: Charles C. Thomas; 1974.
- Hammen C. Generation of stress in the course of unipolar depression. *Journal of Abnormal Psychology* 1991;100:555–561. [PubMed: 1757669]
- Hammen C. Stress and depression. *Annual Review of Clinical Psychology* 2005;1:293–319.

- Hobfoll SE. Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist* 1989;44:513–524. [PubMed: 2648906]
- Holmes TH, Rahe RH. The social readjustment rating scale. *Journal of Psychosomatic Research* 1967;11:213–218. [PubMed: 6059863]
- Hudgens, RW. Personal catastrophe and depression: A consideration of the subject with respect to mentally ill adolescents, and a requiem for retrospective life-event studies. In: Dohrenwend, BS.; Dohrenwend, BP., editors. *Stressful life events: Their nature and effects*. New York: Wiley and Sons; 1974. p. 119-134.
- Jenkins CD, Hurst MW, Rose RM. Life changes: Do people really remember? *Archives of General Psychiatry* 1979;36:379–384. [PubMed: 426603]
- Johnson, JH.; Bradlyn, AS. Assessing stressful life events in childhood and adolescence. In: Karoly, P., editor. *Handbook of child health assessment: Bio perspectives*. New York: Wiley; 1988. p. 303-331.
- Kanner AD, Coyne JC, Schaefer C, Lazarus RS. Comparison of two modes of stress measurement: Daily hassles and uplifts versus major life events. *Journal of Behavioral Medicine* 1981;4:1–39. [PubMed: 7288876]
- Kasl, SV.; Rodriguez, E.; Lasch, KE. The impact of unemployment on health and well-being. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 111-131.
- Katschnig, H. Measuring life stress—a comparison of the checklist and the panel technique. In: Katschnig, H., editor. *Life events and psychiatric disorders*. New York: Cambridge University Press; 1986. p. 74-106.
- Kendler, KS. Adversity, stress, and psychopathology. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 477-485.
- Kendler KS, Hettema JM, Butera F, Gardner CO, Prescott CA. Life event dimensions of loss, humiliation, entrapment, and danger in the prediction of onsets of major depression and generalized anxiety. *Archives of General Psychiatry* 2003;60:789–796. [PubMed: 12912762]
- Kendler KS, Kessler RC, Walters EE, MacLean C, Neale MC, Heath AC, et al. Stressful life events, genetic liability, and onset of an episode of major depression in women. *American Journal of Psychiatry* 1995;152:833–842. [PubMed: 7755111]
- Kessler RC. The effects of stressful life events on depression. *Annual Review of Psychology* 1997;48:191–214.
- Kessler RC, Davis CG, Kendler KS. Childhood adversity and adult psychiatric disorder in the U.S. national comorbidity survey. *Psychological Medicine* 1997;27:1101–1119. [PubMed: 9300515]
- Kessler RC, Sonnega A, Bromet E, Nelson CB. Post-traumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry* 1995;52:1048–1060. [PubMed: 7492257]
- Kihlstrom, JF.; Eich, E.; Sandbrand, D.; Tobias, BA. Emotion and memory: Implications for self-report. In: Stone, AA., et al., editors. *The science of self-report: Implications for research and practice*. Mahwah: NJ: Erlbaum; 2000.
- Kilpatrick, DG.; Resnick, HS.; Saunders, BE.; Best, CL. Rape, other violence against women, and post-traumatic stress disorder: Critical issues in assessing the adversity-stress-psychopathology relationship. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 161-176.
- King DW, King LA, Erickson DJ, Huang MT, Sharkansky EJ, Wolfe J. Posttraumatic stress disorder and retrospectively reported stressor exposure: A longitudinal prediction model. *Journal of Abnormal Psychology* 2000;109:624–633. [PubMed: 11195986]
- Klein DN, Rubovits DR. The reliability of subjects' reports on stressful life events inventories: A longitudinal study. *Journal of Behavioral Medicine* 1987;10:501–512. [PubMed: 3430591]
- Koenen, K.; Stellman, J.; Stellman, S.; Dohrenwend, B. . . Combat exposure reporting and relationship to PTSD in Vietnam veterans; Presentation at the International Society for Traumatic Stress Studies; San Antonio, TX. 2000 Nov.
- Kohn, R.; Dohrenwend, BP.; Mirotznik, J. Adversity, stress and psychopathology. Dohrenwend, BP., editor. New York: Oxford University Press.; 1998. p. 235-284.

- Kubany ES, Haynes SN, Leisen MB, Owens JA, Kaplan AS, Watson SB, et al. Development and preliminary validation of a brief broad-spectrum measure of trauma exposure: The traumatic life events questionnaire. *Psychological Assessment* 2000;12:210–224. [PubMed: 10887767]
- Kulka, RA.; Schlenger, WE.; Fairbank, JA.; Hough, RL.; Jordan, BK.; Marmar, CR., et al. *Trauma and the Vietnam War generation*. New York: Brunner/Mazel; 1990.
- Lazarus, RS.; Folkman, S. *Stress, appraisal and coping*. New York: Springer; 1984.
- Lennon MC, Dohrenwend BP, Zautra AZ, Marbach JJ. Coping and adaption to facial pain in contrast to other stressful life events. *Journal of Personality and Social Psychology* 1990;59:1040–1050. [PubMed: 2266481]
- Lepore, SL. Measurement of chronic stressors. In: Cohen, S., et al., editors. *Measuring stress: A guide for health and social scientists*. New York: Oxford University Press; 1997. p. 102-120.
- Levav I, Krasnoff L, Dohrenwend BS. Israeli PERI life event scale: Ratings of events by a community sample. *Israel Journal of Medical Sciences* 1981;17:176–183. [PubMed: 7228643]
- Lewinsohn PM, Rohde P, Gau JM. Comparability of self-report checklist and interview data in the assessment of stressful life events in young adults. *Psychological Reports* 2003;93:459–471. [PubMed: 14650673]
- Lloyd C. Life events and depressive disorder reviewed. *Archives of General Psychiatry* 1980;37:541–548. [PubMed: 7377910]
- Lyketsos CG, Nestadt G, Cwi J, Heithoff K, Eaton WW. The life chart interview: A standardized method to describe the course of psychopathology. *International Journal of Methods in Psychiatric Research* 1994;4:143–155.
- Mazure CM, Bruce ML, Maciejewski PK, Jacobs SC. Adverse life events and cognitive-personality characteristics in the prediction of major depression and antidepressant response. *American Journal of Psychiatry* 2000;157:896–903. [PubMed: 10831468]
- McQuaid JR, Monroe SM, Roberts JE, Johnson SL, Garamoni G, Kupfer DJ, et al. Toward the standardization of life stress assessment: Definitional discrepancies and inconsistencies in methods. *Stress Medicine* 1992;8:47–56.
- McQuaid JR, Monroe SM, Roberts JE, Kupfer DJ, Frank E. A comparison of two life stress assessment approaches: Prospective prediction of treatment outcome in recurrent depression. *Journal of Abnormal Psychology* 2000;109:787–791. [PubMed: 11196005]
- Mechanic D. Some problems in the measurement of stress and social readjustment. *Journal of Human Stress* 1975;1:43–48. [PubMed: 1235117]
- Meyer, A. The life chart and the obligation of specifying positive data in psychopathological diagnosis. In: Winters, EG., editor. *The collected papers of Adolf Meyer: Vol. 3. Medical teaching*. Baltimore: Johns Hopkins; 1951.
- Miller PM, Salter DP. Is there a short-cut? An investigation into the life event interview. *Acta Psychiatrica Scandinavica* 1984;70:417–427. [PubMed: 6516891]
- Mollica, RF.; Poole, C.; Tor, S. Symptoms, functioning, and health problems in a massively traumatized population: The legacy of the Cambodian tragedy. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 161-176.
- Monroe SM. Assessment of life events. *Archives of General Psychiatry* 1982;32:606–610. [PubMed: 7092492]
- Monroe SM, Harkness KL. Life stress, the “Kindling” hypothesis, and the recurrence of depression: Considerations from a life stress perspective. *Psychological Review* 2005;112:417–445. [PubMed: 15783292]
- Monroe SM, Kupfer DJ, Frank E. Life stress and treatment outcomes: I. Response during index episode. *Journal of Consulting and Clinical Psychology* 1992;60:718–724. [PubMed: 1401387]
- Monroe SM, Roberts JE. Conceptualizing and measuring life stress: Problems. *Stress Medicine* 1990;6:209–216.
- Morgan, CA. . . The evolution of memory for traumatic events; Paper presented at the Annual Meeting of the International Society for Traumatic Stress Studies; Montreal. 1997 Nov.
- Neugebauer R. Reliability of life-event interviews with outpatient schizophrenics. *Archives of General Psychiatry* 1983;40:378–383. [PubMed: 6838318]

- Neugebauer, R. The reliability of life-events reports. In: Dohrenwend, BS.; Dohrenwend, BP., editors. *Stressful life events and their contexts*. New Brunswick, NJ: Rutgers University Press; 1984. p. 85-107.
- Norman RMG, Malla AK. Stressful life events and schizophrenia I: A review of the research. *British Journal of Psychiatry* 1993;162:161–166. [PubMed: 8435685]
- Norris FH. Epidemiology of trauma: Frequency and impact of different potentially traumatic events on different demographic groups. *Journal of Consulting and Clinical Psychology* 1992;3:409–418. [PubMed: 1619095]
- O'Hare WP. Poverty in America: Trends and new patterns. *Population Bulletin* 1989;49:1–45.
- Ormel J, Oldehinkel AJ, Brilman EI. The interplay and etiological continuity of neuroticism, difficulties, and life events in the etiology of major and subsyndromal, first and recurrent depressive episodes in later life. *American Journal of Psychiatry* 2001;158:885–891. [PubMed: 11384895]
- Paykel, ES. Life stress and psychiatric disorder: Applications of the clinical approach. In: Dohrenwend, BS.; Dohrenwend, BP., editors. *Stressful life events: Their nature and effects*. New York: Wiley; 1974. p. 135-149.
- Paykel ES. Methodological aspects of life events research. *Journal of Psychosomatic Research* 1983;27:341–352. [PubMed: 6668560]
- Paykel ES. Methodology of life events research. *Advanced Psychosomatic Medicine* 1987;17:13–29.
- Paykel ES. The interview for recent life events. *Psychological Medicine* 1997;27:301–310. [PubMed: 9089823]
- Paykel ES. The evolution of life events research in psychiatry. *Journal of Affective Disorders* 2001;62:141–149. [PubMed: 11223102]
- Paykel ES, Prusoff BA, Uhlenhuth EH. Scaling of life events. *Archives of General Psychiatry* 1971;25:340–347. [PubMed: 5116988]
- Post RM. Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *American Journal of Psychiatry* 1992;149:999–1010. [PubMed: 1353322]
- Rabkin JG, Struening EL. Life events, stress and illness. *Science* 1976;194:1013–1020. [PubMed: 790570]
- Rahe RH, Arthur RJ. Life change and illness studies: Past history and future directions. *Journal of Human Stress* 1978;4:3–15. [PubMed: 346993]
- Raphael KG, Cloitre M. Does mood-congruence or causal search govern recall bias? A test of life event recall. *Journal of Clinical Epidemiology* 1994;47:555–564. [PubMed: 7730881]
- Raphael KG, Cloitre M, Dohrenwend BP. Problems of recall and misclassification with checklist methods of measuring stressful life events. *Health Psychology* 1991;10:62–74. [PubMed: 2026132]
- Resnick HS, Kilpatrick DG, Dansky BS, Saunders BE, Best CL. Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology* 1993;6:984–991. [PubMed: 8113499]
- Roemer L, Litz B, Orsillo S, Ehlich PJ, Friedman MJ. Increases in retrospective accounts of war-zone exposure over time: The rise of PTSD symptom severity. *Journal of Traumatic Stress* 1998;11:597–605. [PubMed: 9690197]
- Rudolph K, Hammen C. Age and gender as determinants of stress exposure, generation, and reactions in youngsters: A transactional perspective. *Child Development* 1999;70:660–677. [PubMed: 10368914]
- Rutter M. Meyerian psychobiology, personality development, and the role of life experiences. *American Journal of Psychiatry* 1986;143:1077–1087. [PubMed: 3529992]
- Sarason IG, Johnson JH, Siegel JM. Assessing the impact of life changes: Development of the life experiences survey. *Journal of Consulting and Clinical Psychology* 1978;46:932–946. [PubMed: 701572]
- Sarason, IG.; Sarason, BR. Life change, social support, coping and health. In: Kaplan, RM.; MH, Cricqui, editors. *Behavioral epidemiology and disease prevention*. 84. New York: Plenum Press Publishing; 1985. p. 219-236.

- Schaeffer, NC. Asking questions about threatening topics: A selective overview. In: Stone, AA., et al., editors. *The science of self-report: Implications for research and practice*. Mahwah, NJ: Erlbaum; 2000. p. 105-121.
- Schless AP, Schwartz L, Goetz C, Mendels J. How depressives view the significance of life events. *British Journal of Psychiatry* 1974;125:406-410. [PubMed: 4425778]
- Shrout P, Link BG, Dohrenwend BP, Skodol AE, Stueve A, Mirotznik G. Characterizing life events as risk factors for depression: The role of fateful loss events. *Journal of Abnormal Psychology* 1989;98:460-467. [PubMed: 2592681]
- Southwick SM, Morgan CA III, Nicolau AL, Charney DS. Consistency of memory for combat-related traumatic events in veterans of operation desert storm. *American Journal of Psychiatry* 1997;154:173-177. [PubMed: 9016264]
- Steele GP, Henderson S, Duncan-Jones P. The reliability of reporting adverse experiences. *Psychological Medicine* 1980;10:301-306. [PubMed: 7384331]
- Stueve, A.; Dohrenwend, BP.; Skodol, AE. Relationships between stressful life events and episodes of major depression and nonaffective psychotic disorders: Selected results from a New York risk factor study. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 341-357.
- Tennant C, Andrews G. The pathogenic quality of life event stress in neurotic impairment. *Archives of General Psychiatry* 1978;35:859-863. [PubMed: 678039]
- Tennant C, Bebbington P, Hurry J. The role of life events in depressive illness: Is there a substantial causal relation? *Psychological Medicine* 1981;11:370-389.
- Theorell, T. Life events before and after the onset of a premature myocardial infarction. In: Dohrenwend, BS.; Dohrenwend, BP., editors. *Stressful Life Events*. New York: Wiley; 1974. p. 101-117.
- Thoits, PA. Dimensions of life events that influence psychological distress: An evaluation and synthesis of the literature. In: Kaplan, HB., editor. *Psychosocial stress*. New York: Academic Press, Inc; 1983. p. 33-303.
- Thoits PA. Stress, coping, and social support processes: Where are we? What next? *Journal of Health and Social Behavior* 1995;5:53-79. [PubMed: 7560850]
- Tourangeau, R. Remembering what happened: Memory errors and survey reports. In: Stone, AA., et al., editors. *The science of self-report: Implications for research and practice*. Mahwah, NJ: Erlbaum; 2000. p. 29-47.
- Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science* 1998;280:867-873. [PubMed: 9572724]
- Turner RJ, Lloyd DA. Lifetime trauma and mental health: The significance of cumulative adversity. *Journal of Health and Social Behavior* 1995;36:360-376. [PubMed: 8719054]
- Turner RJ, Lloyd DA. Stress burden and the lifetime incidence of psychiatric disorder in young adults: Racial and ethnic contrasts. *Archives of General Psychiatry* 2004;61:481-488. [PubMed: 15123493]
- Turner, RJ.; Wheaton, B. Checklist measurement of stressful life events. In: Cohen, S., et al., editors. *Measuring stress: A guide for health and social scientists*. New York: Oxford University Press; 1997. p. 29-58.
- Vinokur A, Selzer ML. Desirable versus undesirable life events: Their relationship to stress and mental distress. *Journal of Personality and Social Psychology* 1975;32:329-337. [PubMed: 1239500]
- Weitz, H. Direct examination of lay witness. In: Goldberg, NA.; Joseph, GP., editors. *Preparing for and trying the civil lawsuit*. New York: New York State Bar Association; 1987. p. 189-219.
- Wethington, E.; Brown, GW.; Kessler, RC. Interview measurement of stressful events. In: Cohen, S., et al., editors. *Measuring stress: A guide for health and social scientists*. New York: Oxford University Press; 1997. p. 59-79.
- Widom, CS. Childhood victimization: Early adversity and subsequent psychopathology. In: Dohrenwend, BP., editor. *Adversity, stress and psychopathology*. New York: Oxford University Press; 1998. p. 81-95.
- Williams CL, Uchiyama C. Assessment of life events during adolescence: The use of self-report inventories. *Adolescence* 1989;14:95-118. [PubMed: 2658486]

- Wittchen HU, Essau CA, Hecht H, Teder W, Pfister H. Reliability of life event assessments: Test-retest reliability and fall-off effects of the Munich Interview for the Assessment of Life Events and conditions. *Journal of Affective Disorders* 1989;16:77–91. [PubMed: 2521655]
- Yager J, Grant I, Sweetwood HL, Gerst M. Life event reports by psychiatric patients, non-patients, and their partners. *Archives of General Psychiatry* 1981;38:343–347. [PubMed: 7212966]
- Zautra AJ, Guarnaccia CA, Dohrenwend BP. Measuring small life events. *American Journal of Community Psychology* 1986;14:629–655. [PubMed: 3799554]
- Zautra AJ, Reich JW. Life events and perceptions of life quality: Developments in a two-factor approach. *Journal of Community Psychology* 1983;11:121–132.
- Zimmerman M, O'Hara MW, Corenthal CP. Symptom contamination of life event scales. *Health Psychology* 1984;3:77–81. [PubMed: 6536484]
- Zimmerman M, Pfohl B, Stangl D. Life events assessment of depressed patients: A comparison of self-report and interview formats. *Journal of Human Stress* 1986 Spring::13–19. [PubMed: 3559183]
- Zuckerman L, Oliver JM, Hollingsworth HH, Austrin HR. A comparison of life events scoring methods as predictors of psychological symptomatology. *Journal of Human Stress* 1986:64–70. [PubMed: 3559190]

Acknowledgements

This work and some of the studies on which it draws have been supported by Grants K05MH14663, MH26208, and MH59627 from the National Institute of Mental Health. I thank Benjamin Adams, Catherine Douglass, Stephani Hatch, Karestan Koenen, Itzhak Levav, and J. Blake Turner for valuable comments and criticisms.

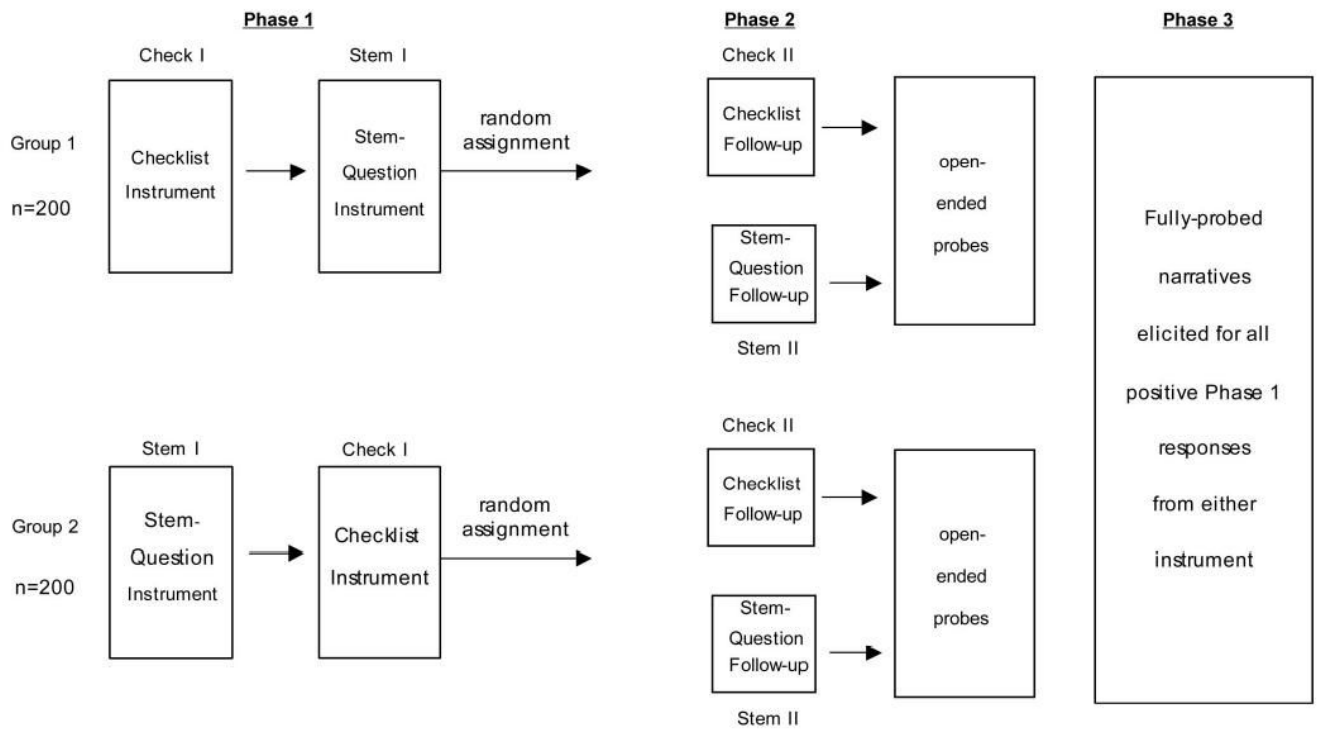


Figure 1. Three-phase design for comparing reliability and validity of checklist and stem question screening instruments for measuring major stressful events.

Table 1**SEPRATE Magnitude Rating**

Magnitude is measured in terms of the amount of change in usual activity lasting at least a week that most people (i.e., average person) would undergo as a consequence of an event. Usual activities are those that characterize family relations, work, and nonwork-related social and leisure activities. Change consists of new things done, things no longer done, or both in these domains of activity. These ratings are made by the following graded categories (note that changes, however large, lasting less than a week are scored 5).

- 1 An extremely large amount of change lasting week or more (e.g., A woman, pregnant with her first child, experiences sudden death of her husband.)
 - 2 A large amount of change lasting a week or more (e.g., A man's wife gave birth to a stillborn child who would have been their first and whom both wanted.)
 - 3 A moderate amount of change lasting a week or more (e.g., A woman's husband suffered a bad break of his arm, was unable to work, and had to stay home for a month.)
 - 4 A little change lasting a week or more (e.g., A man completed one job and started another that required application of the same skills.)
 - 5 No change lasting a week or more (e.g., A woman's husband's uncle whom they rarely saw or kept in touch with died.)
-

Note. These ratings are from Dohrenwend, B. P., Raphael, K. G., Schwartz, S., Stueve, A. & Skodol, A. E., "The structured event probe and narrative rating method (SEPRATE) for measuring stressful life events." In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects*, 2nd edition, p. 198. Copyright 1993 by The Free Press, a Division of Simon & Schuster Adult Publishing Group. Adapted with permission of the publisher. All rights reserved.

Variability Within Event Categories on Rated Amount of Change Based on Narrative Details of What Actually Occurred

Table 2

Life Change Unit (LCU) scores of Event Categories compared with "Marriage" with a modular score of 500	% Large or extremely large amount of change	% Moderate amount of change	% Little change	% No change	n
Serious physical illness started or got worse—668	47.2	27.8	8.3	16.7	(36)
Relationship with spouse or mate changed for the worse without separation or divorce—526	41.2	47.1	0.0	11.8	(17)
Family member other than spouse or child died—457	8.3	8.3	29.2	54.2	(72)
Close friend died—457	5.3	15.8	29.8	49.1	(57)
Suffered a financial loss or loss of property not related to work—446	16.3	44.2	18.6	20.9	(43)
Assaulted—383	18.5	22.2	40.7	18.5	(27)
Broke up with a friend—328	0.0	26.1	37.0	37.0	(46)
Laid off—325	13.3	63.3	13.3	10.0	(30)
Had trouble with a boss—322	17.5	35.0	32.5	15.0	(40)
Got involved in a court case—304	9.5	9.5	28.6	52.4	(21)

Note. The data are from Table 1 in "Measuring Life Events: The Problem of Variability Within Event Categories," by B. P. Dohrenwend, B. G. Link, R. Kern, P. E. Shrout, and J. Markowitz, 1990, *Stress Medicine*, 6, 182. Copyright 1990 by John Wiley & Sons, Limited. Reproduced with permission.

* The higher the LCU score, the greater the change.

Table 3
Examples of Checklist Events on Desert Storm Trauma Questionnaire

Item	Percent change (%)
Extreme threat to your personal safety	35.6
Bizarre disfigurement of bodies as a result of wounds	33.9
Seeing others killed or wounded	27.4
Being stationed close to enemy lines	16.9

Note. The data are from Table 2 in "Consistency of Memory for Combat-Related Traumatic Events in Veterans of Operation Desert Storm," *American Journal of Psychiatry*, 154:2, February 1997 by S. M. Southwick, CA. Morgan, A. L. Nicolaou, and D. S. Charney, 1997, p. 175. Reprinted with permission from the *American Journal of Psychiatry*, Copyright 1997, American Psychiatric Association.

Table 4
Ratings of Source as a Basis for Determining Fatefulness

Source	Prelude	Occurrence
0—Completely determined by external circumstance		X
1—Mostly determined by external circumstances		
2—Equally determined by R's behavior and external circumstances		
3—Mostly determined by R's behavior	X	
4—Completely determined by R's behavior		

Note. These ratings are from Dohrenwend, B.P., Raphael, K.G., Schwartz, S., Stueve, A. & Skodol, A.E., "The structured event probe and narrative rating method (SEPRATE) for measuring stressful life events." In L. Goldberger & S. Breznitz (Eds.), *Handbook of stress: Theoretical and clinical aspects, 2nd edition*, p. 198. Copyright 1993 by The Free Press, a Division of Simon & Schuster Adult Publishing Group. Adapted with permission of the publisher. All rights reserved.

Table 5

Comparison of Three Fateful-Loss Exposure Variables in Patient Cases with Major Depression and Community Controls

Index	All incidents considered without restrictions	Incidents verified to be fateful	Incidents verified to be fateful and disruptive
% of subjects who experienced 1 or more of 12 events meeting conditions			
Case	47.9	39.6	27.3
Community	35.4	30.4	9.7
Unadjusted odds ratio	1.68**	1.50*	3.38***
Odds ratio adjusted for demographic variables	1.77**	1.56*	3.22***
Odds ratio adjusted for demographic variables and demoralization	1.29	1.21	2.51**

Note. From Table 2 in "Characterizing life events as risk factors for depression: The role of fateful loss events" by P. E. Shrout, B. G. Link, B. P. Dohrenwend, A. E. Skodol, A. Stueve, & G. Mirotnik. *Journal of Abnormal Psychology*, 98, p. 465. Copyright 1989, American Psychological Association. Adapted with permission of P. E. Shrout and the American Psychological Association.

* $p < .10$.

** $p < .05$.

*** $p < .001$.