The Stressful Life Events Schedule for children and adolescents: development and validation

Douglas E. Williamson*, Boris Birmaher, Neal D. Ryan, Tiffany P. Shiffrin, Jennifer A. Lusky, Julie Protopapa, Ronald E. Dahl, David A. Brent

Department of Psychiatry, University of Pittsburgh School of Medicine, Room E-723, Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Pittsburgh, PA 15213, USA

Received 25 October 2002; received in revised form 15 May 2003; accepted 16 May 2003

Abstract

The development and psychometric properties of the Stressful Life Events Schedule (SLES), an interview instrument to assess stressors in children and adolescents, are described. Children (≤12 years) and adolescents (>12 years) with psychopathology (n = 30) and non-psychiatric controls (n = 30) were interviewed with the SLES about the occurrence of stressful life events during the prior year. To examine concurrent validity of the SLES, all subjects also completed the self-report Life Events Checklist (LEC) and half the sample completed the Life Events and Difficulties Schedule (LEDS). To examine informant validity, the parent/guardian also reported on stressful life events occurring to their child during the year prior to interview. Test–retest reliability of the SLES was examined by re-assessing all children approximately one week after the initial interview. Subjects rated subjectively how stressful an event was on a 4-point scale. Additionally, panel ratings of objective stress and behavior-dependence/independence were made on 4-point scales. The SLES was found to have substantial inter-rater consensus reliability for objective threat (κ = 0.67) and almost perfect reliability for event behavior-dependence/independence (κ = 0.84). Similarly, the test–retest reliability of the SLES was also found to be substantial at the level of specific event comparisons (κ = 0.68). The SLES discriminated between children with and without psychopathology on all measures of stressful life events. Total stressful life events assessed with the SLES concurred well with those assessed by the LEC (ICC = 0.83) and the LEDS (κ = 0.77) although, as expected, examination of specific events showed much smaller overlap between the SLES and the LEC (κ = 0.26). Child–parent agreement for the occurrence of severe events was substantial (κ = 0.73) but tended to be only moderate when all events were examined (κ = 0.48). The results of this study indicate that the SLES has good psychometric properties. The SLES is a useful, cost-effective tool for assessing stressful life events in children and adolescents.

Keywords: Stress; Measurement; Interviews

*Corresponding author. Tel.: +1-412-624-4526; fax: +1-412-624-0223.
E-mail address: williamsonde@msx.upmc.pitt.edu (D.E. Williamson).
1. Introduction

Beginning with Selye’s efforts to describe the general adaptation syndrome stress response (Selye, 1956), there has been considerable interest in quantifying stress and examining its relation to diseases in general (Goldberger and Beznitz, 1993). Research conducted during the past few decades has shown stressful life events to be etiologically related to the onset of depression in children and adolescents (Compas et al., 1994; Goodyer, 1994; Williamson et al., 1998). Moreover, several studies have focused on specific components of the stressful environment, such as events occurring as a result of a person’s involvement (Kessler, 1997; Silberg et al., 1999; Williamson et al., 1995a), that appear to be important for the onset of depression. While stressful life events research in children and adolescents has evolved to the point of identifying stressful environments that are etiologically important for the onset of depression across development, research in this area has been slowed to some degree by the available methods to assess stress.

Instrumentation for measuring stress has been developed predominantly along the lines of respondent-based or checklist methods. One such respondent-based instrument, developed for use with children and adolescents, is the Life Event Record (LER) (Coddington, 1972). The LER and variations of the LER, such as the Life Events Checklist (LEC) (Johnson and McCutcheon, 1980), have been widely used in studies of adverse events among children and adolescents with various psychiatric disorders. For example, stressful life events and symptoms of depression have been shown to be positively correlated (Banez and Compas, 1990; Johnson and McCutcheon, 1980; Mullins et al., 1985; Swearingen and Cohen, 1985), and studies of clinically depressed youth have found stressful life events to be elevated in depressed children (Kashani et al., 1990) and adolescents (Daniels and Moos, 1990; Williamson et al., 1995a). While these studies suggest that depression and stressful life events are significantly related to one another, it cannot be inferred that stressful life events cause depression since they are just as likely to be caused by the depression (Williamson et al., 1998).

There are several limitations to assessing stressful life events with respondent-based instruments (see Katschnig, 1986, for a review). First, since the checklist contains a finite number of events, stressful events may have occurred but not be listed in the checklist and, therefore, not be counted. Second, checklists such as the LER do not differentiate between acute stressors and ongoing or chronic stressors. Third, while stressful life events may pose various levels of adversity for different individuals, all events endorsed are assumed to carry a similar level of adversity for each person. Fourth, whether an event is dependent on (e.g. interpersonal/relationship disturbances) or independent (e.g. death of a relative) of the subject’s behavior is not assessed. Finally, adolescents and/or their parent/guardians are only asked to indicate if the event occurred during some preceding period of time, typically the past year. Thus, the occurrence of each event is not precisely dated, and the exact timing between event occurrence and onset of the depressive episode cannot be determined.

The study of stressful life events among adults has similarly been hindered by the available methods to assess stressful life events. To circumvent these limitations, Brown and colleagues (Brown et al., 1973) pioneered a methodological advancement in the assessment of stressful life events for use with adults using an investigator-based instrument called the Bedford College Life Event and Difficulty Schedule (LEDS) (Brown and Harris, 1978, 1989). This life event interview collects detailed information about the event itself, the timing of its occurrence, and relevant contextual information (details of what the event actually was) for each event. Based on the contextual information, the threat for each event is rated by a panel of raters using standardized rating procedures drawn from hundreds of examples. The LEDS provides a number of methodological advances over life events checklists that include the following: (1) careful dating of event occurrence; (2) a scaling of the degree of stress (“threat”) for each event; (3) the use of ‘event dictionaries’ and consensus judgments to ensure reliability for including life occurrences as events and for accurately coding threat level; (4) ratings
of event-behavior and event-illness independence; (5) scoring of event dimensions (e.g. loss, danger); and (6) definition and inclusion of chronic stressing or difficulties. Several studies have supported the reliability (e.g. inter-rater) and validity (e.g. multiple informants) of the LEDS with adults exhibiting a variety of psychiatric symptomatology (see Brown and Harris, 1978, 1989).

The LEDS was originally adapted for use with adolescents living in the United Kingdom (Monck and Dobbs, 1985). Recently, we have adapted the LEDS for use with children and adolescents living in the United States and have reported that depressed adolescents have significantly more LEDS-defined severe events (those rated at the highest two severity points) prior to the onset of a depressive episode compared to normal control children (Williamson et al., 1998). Our results regarding the potential etiologic significance of stressful life events coincide with the findings from at least one other research group, which were also based on an investigator-based instrument for assessing stressful life events (Goodyer et al., 1985, 1988).

While the LEDS approach represents a significant improvement over respondent-based methods for assessing stressful life events, it is a very expensive and labor-intensive method that involves upwards of 30 h to interview, rate and reach consensus for a single subject. Thus, it is virtually cost prohibitive to use the LEDS in large-scale studies. Moreover, in our pilot work, we found that less than 10% of all events assessed by the LEDS are severe events, those events shown to be most important for the onset of depression in children and adolescents as well as adults (Brown and Harris, 1978, 1989; Williamson et al., 1998). Recently, we compared stressful life events ascertained by the LEDS with those ascertained by the LEC and found the two to have similar discriminant validity (Duggal et al., 2000), though the two instruments appeared to be reporting on different events. While the LEC and LEDS were found to discriminate between depressed and normal controls equally well, our comparisons were limited since the LEC does not code the specific date of an event and does not cover the same range of events covered in the LEDS (Duggal et al., 2000).

To address the weaknesses of existing instruments, we set out to develop an instrument to ascertain stressful life events that incorporated the strengths of the investigator-based LEDS and improved upon the weaknesses of the respondent-based checklists. This report details the development of a new assessment instrument, the Stressful Life Events Schedule (SLES) for use with children and adolescents, and reports on the results of a study designed to examine its validity and reliability. The specific aims of the study were: (1) to examine inter-rater consensus reliability of the SLES; (2) to determine the test–retest reliability of the SLES; (3) to determine if events reported on the SLES significantly discriminated between youth with psychopathology and normal controls; (4) to establish the concurrent validity of the SLES with both the LEC and the LEDS; and finally (5) to compare level of agreement between child and parent reports of stressful life events about events occurring to the child.

2. Methods

2.1. Sample

The sample consisted of 30 children and adolescents with a current psychiatric illness and 30 normal controls. The psychiatrically ill subjects were recruited from the outpatient clinic at Western Psychiatric Institute and Clinic (WPIC) and were required to be in a current episode of a psychiatric disorder. Primary diagnoses within the psychiatrically ill sample included the following: major depressive disorder \( n = 13 \); anxiety \( n = 11 \); depression not otherwise specified \( n = 2 \); and adjustment disorder with depressed \( n = 2 \) or anxious \( n = 2 \) mood. Since the primary focus of the study was to examine the psychometric properties of the SLES, children with a current psychiatric disorder (primarily mood disorders) were included to inflate the base rate of ascertainable events in the sample and to examine the discriminant validity of the SLES. Normal control youngsters were recruited from ongoing studies being conducted at WPIC and were required to be free of any lifetime psychiatric disorders. The current and lifetime psychiatric status of all subjects was assessed with
the Kiddie Schedule for Affective Disorders and Schizophrenia for school-age children, present and lifetime version (kappa values for all diagnoses ≥ 0.80) (Kaufman et al., 1997). For all subjects, 52% were female, 87% were Caucasian, and the average age was 13.0 ± 1.6 years. There were no differences between psychiatrically ill and normal control subjects on any of the demographic characteristics. All subjects had parental consent to participate in the study, and children and adolescents assented to participate. Subjects and their parents were paid $20.00 for each SLES interview and subjects completing the LEDS were paid $50.00.

2.2. Instruments

2.2.1. Life Events Checklist (LEC)

The LEC is a modification of the Life Event Record (LER) (Coddington, 1972) and contains 46 life events plus space for additional life events to be added by the respondent (Johnson and McCutcheon, 1980). The LEC deviates from the LER in that the respondent is asked to indicate if the event was positive or negative and the degree to which the event was stressful or unpleasant (coded on a 4-point scale ranging from no effect to great effect). Units of measure yielded by the LEC include the total number of positive and negative events as well as a weighted total impact of positive and negative events. Test–retest reliability of the LEC has been reported to be 0.69 (positive events) and 0.72 (negative events) for simple unit weights, and 0.71 (positive events) and 0.55 (negative events) using weighted impact scores (Brand and Johnson, 1982).

2.2.2. Life Events and Difficulties Schedule (LEDS)

A version of the LEDS adapted for use with children and adolescents was used (Monck and Dobbs, 1985; Williamson et al., 1998). The LEDS is an investigator-based semi-structured interview schedule designed to elicit relevant contextual information for events and difficulties covering 10 domains: education, work, reproduction, money/possessions, housing, crime/legal, health, romantic relationships, other relationships and miscellaneous events (including deaths). The degree of stress for each event was rated on a 4-point scale (1—marked, 2—moderate, 3—some, 4—little or none) for both short-term (peak threat of the event in the 48 h following the event) and long-term threat (peak threat in the 10–14-day period after the event occurred). Difficulties, lasting a minimum of one month (i.e. 28 days), were rated on a 7-point scale (1—high marked, 2—low marked, 3—high moderate, 4—low moderate, 5—mild, 6—very mild, 7—not/no longer a difficulty). The short- and long-term threat ratings for events and the threat ratings for difficulties were made by comparing the contextual information for each event with a dictionary containing contextual examples based on interviews with adolescents and adults. For each event, the focus of the event was determined to be specifically on the subject alone, on the subject jointly with another person, or on another person, or involving a possession or a pet. After the interviewers (JAL and JP) rated each occurrence (i.e. a chronological composite of events and incidents followed by chronic difficulties), the LEDS chart for each subject was presented to a panel of consensus raters consisting of two raters blind to the adolescent’s diagnostic status. During the consensus meeting, the interviewer presented an overview of the adolescent’s biographical circumstances that included personal and family demographics as well as whom the adolescent’s close relatives and confidants were (since events occurring to them may be eligible for inclusion depending on the nature of the event).

2.2.3. Stressful Life Events Schedule (SLES)

The development of the SLES grew out of our interest in examining the relations between stress and depression in children and adolescents, based on our experience using life events checklists (Birmaher et al., 1994; Williamson et al., 1995a,b, 1996) and interview instruments (Williamson et al., 1998). We sought to create an interview instrument that incorporated the strengths of the LEDS but took a fraction of the time to learn and administer. To achieve this goal, a streamlined approach was taken that involved directly asking if an event occurred and developing anchor ratings that contained specific criteria for rating each event.
occurrence. Toward this end, child, adolescent and parent versions of the SLES were created to (1) probe for the occurrence of an event; (2) use memory aids to date the occurrence of the event; (3) record the duration of the event in order to differentiate between acute and ongoing events; (4) assess the subjective threat of each event; (5) assess the objective threat of each event; (6) assess the behavior-dependence/independence of each event; and (7) denote the focus of each event paralleling the LEDS classification of self, joint, other or pet/possession classifications. The initial development of the SLES is detailed below.

The first step in creating the SLES involved identifying all acute and ongoing stressful life events, in particular those events capable of being rated severe as assessed by the adult and adolescent version of the LEDS (see Brown and Harris, 1989; Williamson et al., 1998). After extracting the events from the LEDS dictionaries, relevant probes for each stressful life event were identified if they were explicitly or implicitly described in either the interview schedule or the dictionary or logically flowed from the question itself.

The second step in creating the SLES involved reviewing all existing life events instruments and adding those potentially negative events not covered by the LEDS as described in the first step above. Self-report measures of life events such as the LER and LEC as well as interview schedules such as the Psychosocial Assessment of Childhood Experiences (PACE) (Sandberg et al., 1993) and other instruments (e.g. Masten et al., 1994; Goodyer et al., 1985) were reviewed and events extracted to be included in the pool of events. Again, only events rated as potentially negative were included in the final version of the instrument. Positive events (e.g. ‘outstanding achievement’) appear not to be related to the onset and presence of psychopathology (e.g. Goodyer, 1994) and were therefore excluded. The final parent and adolescent versions of the SLES contain 77 events (61 in the child version) and allow for the inclusion of additional events not explicitly covered. Events occurring to the child or adolescent and important others (e.g. parents, relatives, friends) are categorized into domains covering education, work, money, housing, crime, health, deaths, romantic relationships and other relationships.

The third step in the development of the SLES involved incorporating aspects of each event to be rated as suggested by current research. As a result of this step, the behavior-dependence/independence rating was included for each event. Research examining stressful environments associated with MDD has shown that the risk for stressful life events is, in part, genetically mediated (Kendler et al., 1993). In turn, active involvement in creating environmental stresses has been shown to increase the risk for depression in adults (e.g. Hammen et al., 1991; Kendler et al., 1995) and adolescents (Daley et al., 1997). Moreover, depressed adolescents exposed to behavior-dependent events (e.g. arguments with family members, breaking up with a boyfriend/girlfriend) have been shown to differentially manifest various depressive symptoms during their episode (Daley et al., 1997; Williamson et al., 1995a). For example, depressed adolescents exposed to a behavior-dependent event were found to have fewer negative-cognitions, less anxiety symptoms and lower depression severity scores (Williamson et al., 1995a) compared with depressed adolescents without a behavior-dependent event. Accordingly, ratings to determine the behavior independence/dependence of each event were included.

The fourth step in the development of the SLES involved examining how well event descriptions generated by LEDS interviews could be classified and rated within the SLES. This step involved randomly selecting descriptions for 300 events rated via the LEDS as part of our pilot project (Williamson et al., 1998) and having the objective threat rating for each separately rated by a panel of raters. The results from this step suggested that the severity of ratings of events between the LEDS and the objective threat ratings given by the SLES were quite comparable with 85% of the 300 events rated exactly the same. As a result of having four raters review the same 300 events and separately rate them, the SLES was edited for inconsistencies and ambiguities. In addition, events mistakenly included or omitted were either removed or added. After revising the SLES, another random sample of 300 event descriptions were chosen and re-rated.
by the same panel of four raters. Again, the general level of agreement between the LEDS and the SLES was excellent with the objective threat ratings found to be identical between the two, 90% of the time. At these initial development steps, detailed rating information was not collected for each event. Therefore, more formal estimates of the convergent validity between the two instruments could not be made.

As a result of the above steps, the initial development of the SLES was complete and the next step was to formally examine its psychometric properties.

2.3. Study design and administration of the assessments

The following study design was used to examine the validity of the SLES. All children and adolescents were initially interviewed face-to-face with the SLES about the occurrence of stressful life events during the year prior to interview. The current interview format of the SLES involves systematically asking the informant if each of the events occurred and, if they did, several follow-up questions are asked for each event to probe for the relevant contextual information needed to include and rate the event. For example, in the housing section, informants are asked if they changed residences during the previous year. If they had, the following questions were asked: Why did you move? Did you want to move? Did you have to change schools? Do you like your new home? Did you feel too far away from friends at this new home? Would you rather live somewhere else? Ongoing stressors that were present during the previous year but had begun outside the interview period were dated back to when they began. To control for potential memory recall bias, half of the sample was randomly chosen to complete the LEC prior to being interviewed with the SLES and the other half of the sample completed the LEC after being interviewed with the SLES. After their child was interviewed, all parents were interviewed face-to-face with the SLES about stressful life events occurring in their child’s life during the past year by the same interviewer. In order to examine the convergent validity between the SLES and the LEDS, 30 youngsters (15 with psychopathology and 15 normal controls) were randomly selected within each of the groups (i.e. normal controls vs. psychopathological subjects) and interviewed with the LEDS after both the subject and his/her parent had been interviewed with the SLES. The psychiatrically ill subjects had the following diagnoses: major depressive disorder \( n=6 \); anxiety \( n=5 \); depression not otherwise specified \( n=1 \); and adjustment disorder with depressed \( n=1 \) or anxious \( n=2 \) mood. The LEDS was systematically administered last in order to weight any potential memory recall biases in favor of the LEDS, thereby increasing the chances that it would perform better compared with the SLES. It was anticipated that initial exposure to the LEC and the SLES might prime respondents to remember more clearly events that had occurred during the prior year as well as increase the accuracy with which they dated the events. The final step in the study involved re-interviewing all children and adolescents with the SLES. On average, the second interview was conducted \( 9.3 \pm 2.6 \) days (range 5–15 days) after the initial SLES interview. All but one subject participated in the SLES retest, and all retest interviews were conducted by telephone.

2.4. Data analysis

All statistical comparisons were made using two-tailed tests with \( \alpha \) set at 0.05. Chi-square analyses were used for categorical variables unless the minimum expected cell size was below 5, in which case Fisher’s exact tests were used. Analysis of variance and analysis of variance with covariance were used when examining continuous outcome measures and exploring the potential effects of covariates. In analyses examining reliability and validity, the intra-class correlation coefficient (ICC) was used for continuous variables and the kappa statistic (\( \kappa \)) was used for categorical variables. All ICC and \( \kappa \) values are reported with their corresponding 95% confidence intervals (CI).

The terminology suggested by Landis and Koch (1977) was adapted to qualitatively describe the level of observer and instrument agreement as originally recommended for the kappa statistic and
was also used for the ICCs. The following classification qualifications were used: <0.00—‘Poor’; 0.0–0.20—‘Slight’; 0.21–0.40—‘Fair’; 0.41–0.60—‘Moderate’; 0.61–0.80—‘Substantial’; and 0.81–1.0—‘Almost Perfect’ (Landis and Koch, 1977).

3. Results

3.1. General description of events ascertained by the SLES

As per the format of the administration of the SLES, there are two steps involved in event ascertainment. The first step involves interviewing the youth and recording whether he/she reports that an event has occurred and obtaining the date, duration and subjective threat level of the event. Based on the interviews conducted for this report, the estimated interviewing time was 39.6 ± 32.9 min (range 15–120 min) and took significantly longer for the psychiatrically ill subjects (55.5 ± 26.7 min, range 15–120 min) compared to the controls (35.5 ± 18.9 min, range 15–120 min) (P < 0.0001). The second step involves the interviewer preparing event descriptions that are presented to a panel of raters. The panel arrives at consensus as to whether an event meets criteria to be included and, if included, the panel rates the objective threat, behavior-dependence and focus of each event. Based on the interviews conducted as part of this study, the estimated chart preparation time is 55.1 ± 38.5 min (range 3–180 min) and took significantly longer in the psychiatrically ill subjects (79.2 ± 46.1 min, range 3–180 min) compared with the controls (48.4 ± 32.1 min, range 5–180 min) (P < 0.0001). The longer interview and chart preparation times for the psychiatrically ill subjects reflected the increased number of events occurring during the prior year relative to the controls (see Section 3.4 below for details).

At the initial interview step, the 60 subjects reported a total of 436 total events. After being reviewed by the consensus panel, 46 events were determined not to meet criteria to be included as a stressful life event. Examples of events excluded were the following: the toilet in the subject’s house broke, the subject wanted to ‘go out’ with a girl but she had a boyfriend, a friend of the subject was in a minor ‘fender bender’ and a friend broke his finger playing basketball. On the subjective threat scale, for the 46 events, 48% were rated ‘little or none’, 33% were rated ‘some’, 13% were rated ‘moderate’ and 6% were rated as ‘high’ stress.

A total of 390 events were rated as being at or above threshold for inclusion and rated by the consensus panel as reported by the child or adolescent. The 20 most frequently reported events among these subjects, accounting for 64% of all events, were as follows: (1) ‘hospitalization/surgery of other’ (n = 27); (2) ‘start of romantic relationship’ (n = 19); (3) ‘school performance problems’ (n = 18); (4) ‘fights/arguments at school’ (n = 17); (5) ‘death of pet’ (n = 16); (6) ‘general health problems of other’ (n = 13); (7) ‘general health problems of subject’ (n = 13); (8) ‘being bullied’ (n = 13); (9) ‘changed schools’ (n = 12); (10) ‘close relative died’ (n = 13); (11) ‘injury/accident of subject’ (n = 11); (12) ‘arguments with siblings’ (n = 10); (13) ‘increased arguments between siblings and parents’ (n = 10); (14) ‘increased arguments with parents’ (n = 10); (15) ‘breakup with boyfriend/girlfriend’ (n = 9); (16) ‘change in physical appearance’ (n = 8); (17) ‘change in parent’s job’ (n = 8); (18) ‘other caught committing crime’ (n = 8); (19) ‘injury/accident of other’ (n = 8); and (20) ‘additional events’ not originally classified in the SLES’ (n = 8).

3.2. Inter-rater consensus reliability

An estimate of the inter-rater consensus reliability for the SLES was achieved by recording all individual ratings at the consensus rating meeting for objective threat, dependence/independence and focus of each event and comparing them. The ratings for six raters, who had rated at least 1000 events, were included. Kappa coefficients were estimated between raters comparing their initial ratings after hearing the event description prior to discussing an event and for their initial rating compared to the final consensus rating. Reliabilities were calculated for exact agreement for each of the 4-point scales and for the dichotomous
Table 1
Inter-rater consensus reliability of the Stressful Life Events Schedule

<table>
<thead>
<tr>
<th></th>
<th>Between raters&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Compared with consensus rating&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Objective threat</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact rating</td>
<td>κ = 0.67</td>
<td>κ = 0.75</td>
</tr>
<tr>
<td>Severe/not severe</td>
<td>κ = 0.70</td>
<td>κ = 0.78</td>
</tr>
<tr>
<td><strong>Behavior-dependence/independence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact rating</td>
<td>κ = 0.84</td>
<td>κ = 0.86</td>
</tr>
<tr>
<td>Dependent/independent</td>
<td>κ = 0.92</td>
<td>κ = 0.94</td>
</tr>
<tr>
<td><strong>Focus of event</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exact rating</td>
<td>κ = 0.93</td>
<td>κ = 0.94</td>
</tr>
</tbody>
</table>

κ = Kappa coefficient.
<sup>a</sup> Ratings for six raters who had rated ≥1000 events were compared with each other prior to consensus ratings.
<sup>b</sup> Ratings for six raters were compared with the final consensus ratings.

outcomes of severe/not severe and behavior-dependent/independent (see Table 1).

For objective threat, the agreement was found to be substantial between raters (κ = 0.67; 95% CI, 0.63–0.71) and between individual raters and the final consensus ratings (κ = 0.75; 95% CI, 0.72–0.79). For dependence/independence ratings, exact levels of agreement were substantial to almost perfect between raters (κ = 0.84; 95% CI, 0.81–0.87), and between raters and the final consensus dependence/independence rating (κ = 0.86; 95% CI, 0.84–0.89). For focus, the level of agreement between raters was almost perfect (κ = 0.93; 95% CI, 0.91–0.95) as was the agreement between raters and the final consensus rating for focus (κ = 0.94; 95% CI, 0.93–0.96). Further analysis of levels of agreement for severe events (rated 3 or 4 on objective threat) showed substantial levels of agreement between raters (κ = 0.70; 95% CI, 0.63–0.77), and between raters and the final consensus rating (κ = 0.78; 95% CI, 0.72–0.84). Analyses examining whether an event was behavior-dependent or not (rated 3 or 4 on dependence scale) showed almost perfect levels of agreement between raters (κ = 0.92; 95% CI, 0.89–0.94) and with the final consensus ratings (κ = 0.94; 95% CI, 0.92–0.96). Age, gender and diagnostic status did not significantly impact any of the reliability estimates for objective threat, behavior-dependence or focus of the event.

3.3. Test–retest reliability of the SLES

A total of 59 out of the sample of 60 youngsters participated in the retest interview phase of the study. The first step in examining the test–retest reliability of the SLES was to compare the total number of events reported at the initial interview with the total number of events reported at the retest interview. For all events, the ICC was 0.93 (95% CI, 0.89–0.96). Stratifying the sample by children (≤12 years) and adolescents (>12 years) yielded ICCs of 0.91 (95% CI, 0.81–0.96) and 0.94 (95% CI, 0.88–0.97), respectively. These same analyses were then repeated using non-severe and severe events separately. Non-severe events, events rated as 1—little/none or 2—some objective threat, had an ICC of 0.90 (95% CI, 0.83–0.94). Normal controls had significantly higher retest reliability compared with children with psychopathology for non-severe events (0.93 vs. 0.83, P < 0.001). Severe events, events rated 3—moderate or 4—high objective threat, had an ICC of 0.70 (95% CI, 0.50–0.82). Adolescents had higher retest reliability compared with children for severe events (0.86 vs. 0.43, P < 0.001) (see Table 2).

To more closely examine the test–retest reliability of the SLES, further analyses were conducted to examine reliability at the event level (see Table 2). For any event, the test–retest reliability was estimated to be κ = 0.68 (95% CI, 0.64–0.72).
Test–retest reliability was higher for females ($\kappa = 0.74$; 95% CI, 0.69–0.79) compared with males ($\kappa = 0.62$; 95% CI, 0.55–0.68) ($\chi^2_{df=2} = 10.74$, $P \leq 0.0047$) for all events. Additionally, test–retest reliability was significantly higher for adolescents ($\kappa = 0.72$; 95% CI, 0.67–0.77) than children ($\kappa = 0.61$; 95% CI, 0.54–0.68) ($\chi^2_{df=2} = 8.55$, $P \leq 0.0139$) for all events. There were no significant differences in test–retest reliability for behavior-dependent/independent events nor was diagnostic group significantly related to test–retest reliability for all events. For non-severe events, the test–retest reliability was estimated to be $\kappa = 0.67$ (95% CI, 0.62–0.71). Test–retest reliability was higher for females ($\kappa = 0.72$; 95% CI, 0.67–0.78) than males ($\kappa = 0.60$; 95% CI, 0.54–0.67) ($\chi^2_{df=2} = 8.75$, $P \leq 0.02$) for non-severe events. There were no significant effects of diagnosis, age, or event dependence/independence for non-severe events. For severe events, the test–retest reliability at the event level was higher than for all events ($\kappa = 0.80$; 95% CI, 0.71–0.89). There were no significant effects of diagnostic group, gender, age, or event dependence/independence on the test–retest reliability estimates for severe events.

### 3.4. Discriminant validity of the SLES

The design of the study included an equal number of youngsters with psychopathology and normal controls. Research has shown that children and adolescents with psychiatric disorders have an elevated rate of stressful life events compared with normal controls (e.g., Williamson et al., 1995a). Accordingly, to determine whether the SLES had discriminant validity, youngsters with psychiatric disorders were compared with normal controls on a range of stressful life event measures.

Subjects with psychiatric disorders had a significantly higher total number of events occur during the year prior to interview compared with normal controls ($8.1 \pm 4.0$ vs. $4.9 \pm 3.5$, $t_{58} = 3.35$, $P \leq 0.001$) (see Table 3). Further analyses examining behavior-independent events ($P \leq 0.03$) and behavior-dependent events ($P \leq 0.005$) similarly showed youth with psychopathology experienced an increased number of these events compared with normal controls. Additional analyses examined the total number of severe events, those events objectively rated as having a 3—moderate or 4—high objective threat level. Once again, youth with psychopathology experienced a significantly increased number of severe events during the past year compared with normal controls ($1.3 \pm 1.8$ vs. $0.3 \pm 0.6$, $t_{58} = 2.77$, $P \leq 0.008$), which remained true for both behavior-independent ($P \leq 0.03$) and behavior-dependent ($P \leq 0.03$) severe events.

In addition to examining the total number of events, the total sum stress levels of both subjectively and objectively assessed threat were examined. Youth with psychopathology experienced significantly more total subjective stress during the past year compared with normal controls ($18.2 \pm 11.0$ vs. $9.9 \pm 7.2$, $t_{58} = 3.47$, $P \leq 0.001$). Similarly, the total subjective stress for both behavior-independent events ($P \leq 0.02$) and behavior-dependent events ($P \leq 0.002$) was significantly greater in the psychiatrically ill youth compared with normal controls. Analyses examining objective stress levels produced the same pattern of results. Psychiatrically ill youth had significantly greater overall objective stress levels during the year prior to interview than normal controls.

### Table 2

Test–retest reliability of the Stressful Life Events Schedule

<table>
<thead>
<tr>
<th></th>
<th>Any events</th>
<th>Specific events</th>
</tr>
</thead>
<tbody>
<tr>
<td>All events</td>
<td>ICC = 0.93</td>
<td>$\kappa = 0.68$</td>
</tr>
<tr>
<td>Non-severe events</td>
<td>ICC = 0.90</td>
<td>$\kappa = 0.67$</td>
</tr>
<tr>
<td>Severe events</td>
<td>ICC = 0.70</td>
<td>$\kappa = 0.80$</td>
</tr>
</tbody>
</table>

$ ICC = $ intra-class correlation coefficient. 
$ \kappa = $ Kappa coefficient.

* Compared total events reported at the initial interview with total number of events reported at the retest interview, 
* Examined specific event occurrence at the initial and retest interviews, 
* Events rated as 1—little/none or 2—some objective threat on a 4-point scale, 
* Events rated as 3—moderate or 4—high objective threat on a 4-point scale, 
* Higher for adolescents than children, 
* Higher for females than males, 
* Higher for control children than psychopathological children.
Table 3
Discriminant validity of the Stressful Life Event Schedule

<table>
<thead>
<tr>
<th></th>
<th>Psychopathology</th>
<th>Normal controls</th>
<th>Statistic</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 30</td>
<td>n = 30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total number of events</td>
<td>8.1 ± 4.0</td>
<td>4.9 ± 3.5</td>
<td>t = 3.35</td>
<td>0.0014</td>
</tr>
<tr>
<td>Behavior-independent</td>
<td>5.0 ± 3.0</td>
<td>3.4 ± 2.4</td>
<td>t = 2.29</td>
<td>0.0256</td>
</tr>
<tr>
<td>Behavior-dependent</td>
<td>3.1 ± 2.5</td>
<td>1.4 ± 1.8</td>
<td>t = 2.96</td>
<td>0.0045</td>
</tr>
<tr>
<td>Total number of severe</td>
<td>1.3 ± 1.8</td>
<td>0.3 ± 0.6</td>
<td>t = 2.77</td>
<td>0.0076</td>
</tr>
<tr>
<td>events</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior-independent</td>
<td>0.7 ± 0.9</td>
<td>0.2 ± 0.6</td>
<td>t = 2.19</td>
<td>0.0325</td>
</tr>
<tr>
<td>Behavior-dependent</td>
<td>0.6 ± 1.2</td>
<td>0.1 ± 0.3</td>
<td>t = 2.29</td>
<td>0.0255</td>
</tr>
<tr>
<td>Overall subjective stress</td>
<td>18.2 ± 11.0</td>
<td>9.9 ± 7.2</td>
<td>t = 3.47</td>
<td>0.0010</td>
</tr>
<tr>
<td>Behavior-independent</td>
<td>11.5 ± 7.8</td>
<td>7.2 ± 5.2</td>
<td>t = 2.49</td>
<td>0.0155</td>
</tr>
<tr>
<td>Behavior-dependent</td>
<td>6.7 ± 5.8</td>
<td>2.7 ± 3.8</td>
<td>t = 3.21</td>
<td>0.0021</td>
</tr>
<tr>
<td>Overall objective stress</td>
<td>14.7 ± 8.8</td>
<td>7.5 ± 5.7</td>
<td>t = 3.79</td>
<td>0.0004</td>
</tr>
<tr>
<td>Behavior-independent</td>
<td>8.8 ± 5.9</td>
<td>5.3 ± 3.8</td>
<td>t = 2.68</td>
<td>0.0095</td>
</tr>
<tr>
<td>Behavior-dependent</td>
<td>5.9 ± 5.4</td>
<td>2.1 ± 2.8</td>
<td>t = 3.42</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

* Total number of events determined by consensus panel to meet criteria for inclusion.
* Total number of events rated as 3—moderate or 4—high on objective threat.
* Sum total of subjectively rated stress.
* Sum total of objectively rated stress.

(14.7 ± 8.8 vs. 7.5 ± 5.7, t = 3.79, P ≤ 0.0004). Similarly, the total objective stress levels were significantly higher in psychiatrically ill youth for behavior-independent (P ≤ 0.02) and behavior-dependent (P ≤ 0.02) events compared with normal controls.

The above analyses were repeated exploring the potential effects of gender, order in which the SLES interview was conducted (i.e. before or after completing the LEC), and age of the subject (child vs. adolescent). For all analyses, neither gender nor order of assessment was significantly related to the occurrence of stressful life events. However, analyses examining age showed that adolescents were significantly more likely than children to have a behavior-dependent event (F1,56 = 12.73, P ≤ 0.0007) and to score higher on the total sum of both objective (F1,56 = 9.90, P ≤ 0.0026) and subjective (F1,56 = 10.32, P ≤ 0.0022) stress (see Table 3). There was no effect of age on any of the behavior-independent stressful life event measures. Further, there was no diagnostic group by age interaction for any of the stressful life event measures.

3.5. Concurrent validity of the SLES with the LEC

For these comparisons, all events originally ascertained with the SLES (n = 441) were included and corresponding subjective threat ratings were used since they were directly comparable to the subjective ratings elicited by the LEC. The first step in examining the concurrent validity of the SLES and the LEC involved examining the total number of events reported by each and computing the intraclass correlation coefficient (ICC). For total number of events, the ICC was estimated at 0.80 (95% CI, 0.67–0.88). For total number of events rated as 1—little/none or 2—some threat on each instrument, the ICC was estimated at 0.64 (95% CI, 0.40–0.78). For total number of events subjectively rated as 3—moderate to 4—high threat on each of the instruments, the ICC was estimated at 0.83 (95% CI, 0.72–0.90). There was no effect of age, gender, diagnostic status, or order of the SLES-LEC administration on the concurrent validity estimates between the two instruments.

One of the limitations posed by the above analysis is the assumption that the events reported by the SLES and the LEC are the same events (Gorman, 1993; Katschnig, 1986). Therefore, events on the LEC were assigned SLES event classification codes and analyses examined the agreement between the two instruments at the event level. At the event level, the SLES and the
Table 4
Concurrent validity of the Stressful Life Events Schedule (SLES) with the Life Event and Difficulties Schedule (LEDS)

<table>
<thead>
<tr>
<th>Subjects with a SLES severe event</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subjects with a LEDS severe event</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>No</td>
<td>3</td>
<td>19</td>
</tr>
</tbody>
</table>

κ = 0.77; 95% CI, 0.53–1.0

LEC were found to have slight concurrent validity for all events (κ = 0.31; 95% CI, 0.26–0.36), fair concurrent validity for non-severe events (κ = 0.47; 95% CI, 0.42–0.52), and moderate concurrent validity for severe events (κ = 0.61; 95% CI, 0.55–0.66). Neither diagnostic group, gender, age, nor the order of the SLES-LEC administration was differentially related to concurrent validity estimates for either all or only severe events.

3.6. Concurrent validity of the SLES with the LEDS

A total of 30 youngsters were also interviewed with the LEDS at the initial interview period. The SLES and the LEDS were found to have substantial concurrent validity for specific severe events (κ = 0.77; 95% CI, 0.53–1.0). All eight subjects with a LEDS-defined severe event had a SLES-defined severe event while three additional subjects had a SLES-defined severe event only (see Table 4).

3.7. Informant variance: child vs. parent reports of stressful life events

For all 60 subjects, both youngsters and their parents were interviewed about stressful life events that occurred to the child during the past year. Based on the parent interviews conducted for this study, interviews took an average of 55.8 ± 27.2 min to complete (range 20–120 min) and 80.2 ± 45.4 min to writeup and prepare for consensus (range 20–180 min). Analyses examining the total number of stressful life events showed substantial to almost perfect agreement between youngsters’ and parents’ reports, ICC = 0.81 (95% CI, 0.68–0.89). Analyses examining non-severe events (ICC = 0.73, 95% CI, 0.54–0.84) and severe events (ICC = 0.75, 95% CI, 0.58–0.85) showed substantial levels of agreement. Neither age, gender, nor diagnostic status was significantly related to the child-parent level of agreement in SLES reports (see Table 5).

Further analyses were conducted examining the occurrence of each event reported by the two informants. For all stressful life events, the child–parent agreement was moderate (κ = 0.48; 95% CI, 0.43–0.52) as was the level of agreement for non-severe events (κ = 0.46; 95% CI, 0.41–0.50). For severe events, the level of agreement was moderate (κ = 0.73; 95% CI, 0.63–0.83) (see Table 5). There were no significant effects of diagnostic group, gender, or age on the agreement between child–parent reports for analyses examining all events or severe events only.

4. Discussion

Scientific inquiry into the etiologic contribution of environmental stress for child and adolescent psychiatric disorders has proliferated during the past few decades. Evolving out of these research efforts, hypotheses regarding the role of specific stressors, or ‘candidate environments’, have emerged. For example, behavior-dependent events, or those events brought about by the child’s or

Table 5
Child vs. parent reports for the Stressful Life Events Schedule

<table>
<thead>
<tr>
<th></th>
<th>Any events</th>
<th>Specific events</th>
</tr>
</thead>
<tbody>
<tr>
<td>All events</td>
<td>ICC = 0.81</td>
<td>κ = 0.48</td>
</tr>
<tr>
<td>Non-severe events</td>
<td>ICC = 0.73</td>
<td>κ = 0.46</td>
</tr>
<tr>
<td>Severe events</td>
<td>ICC = 0.75</td>
<td>κ = 0.73</td>
</tr>
</tbody>
</table>

ICC = intra-class correlation coefficient.
κ = Kappa coefficient.

* Compared total events reported at the initial interview with total number of events reported at the retest interview.

b Examined specific event occurrence at the initial and retest interviews.

c Events rated as 1—little/nothing or 2—some objective threat on a 4-point scale.

d Events rated as 3—moderate or 4—high objective threat on a 4-point scale.
adolescent’s actions, have been shown to be elevated among children of depressed women (Hammen et al., 1991), elevated among depressed adolescents (Williamson et al., 1995a), genetically mediated (Silberg et al., 1999) and predictive of the onset of depression (Goodyer et al., 2000). However, due to the general assessment of stressors in at least some of these studies (e.g. Silberg et al., 1999; Williamson et al., 1995a), a deeper understanding of the nature of behavior-dependent requires further study.

Traditionally, two approaches to measuring life events in children and adolescents have been taken which rely on respondent-based and interviewer-based methods. As reviewed above, the general measurement of environmental stressors assessed in the respondent-based or checklist approach limits the degree to which one can examine the etiologic importance of environmental stressors. At the other end of the spectrum, interviewer-based methods for assessing environmental stressors, for example, the LEDS (Monck and Dobbs, 1985; Williamson et al., 1998) or LEDS-based approaches (Sandberg et al., 1993), require lengthy training periods, many hours to administer, and lengthy periods to rate and then reach consensus. In fact, in our adaptation and application study of the LEDS, some cases took in excess of 30 h to complete in comparison to the SLES, which took an average, of 1 h to complete. Thus, while interviewer-based methods such as the LEDS represent quite a significant improvement over respondent-based methods, their labor-intensive application and related costs have limited their wide-scale use. The SLES takes considerably less time for three main reasons. First, event occurrences are asked about directly in the SLES while the LEDS involves a general interview process and does not simply ask whether an event occurred or not. Second, objective threat ratings for each event have succinct anchoring details that are used to decide if an event is to be coded a 1, 2, 3 or 4 on objective threat. In the LEDS system, dictionaries containing hundreds of event and difficulty examples are used to compare and contrast the context of each event in order to anchor stress ratings. It appeared that it was in this regard that the SLES offered the greatest savings of time. Finally, the LEDS has over 20 subscales that are rated for each event that has evolved out of LEDS-based research conducted in adults (e.g. ‘danger’). Only the objective threat, behavior dependence, and focus of each event are assessed in the SLES.

Based on our research experience examining the relation between stress and depression in youngsters using both respondent-based (Birmaher et al., 1994; Williamson et al., 1995a,b, 1996) and interviewer-based (Duggal et al., 2000; Williamson et al., 1998) assessment methods, we set out to create an interviewer-based life events assessment instrument that significantly improved upon the weaknesses of the respondent-based assessments and yet incorporated the strengths of the investigator-based instruments. The resulting instrument described in this report, the Stressful Life Events Schedule (SLES), takes an average of 1 h to administer and rate and was found to have good concurrent validity with the LEDS ($\kappa = 0.77$). The validity and reliability of the SLES are discussed in more detail below.

4.1. Limitations

Prior to considering the implications of the results of this study, the several limitations of the current report are reviewed. First, all interviews for the same subject (e.g. initial and retest) were administered by the same interviewer. Interviewers were specifically instructed to probe about events relative to the subject, discussing them in each interview separately. However, it is possible that interviewers differentially probed for information based on information from other interviews, thereby inflating the estimates of reliability and validity. It might have been preferable to have different interviewers for each interview. Second, specific details about events reported in the LEC were not gathered after the LEC was completed. Though this is not standard procedure for administering the LEC, the information would have been helpful to specifically identify the discrepancies between the LEC and the SLES. Third, we only assessed inter-rater consensus reliability that was likely influenced by the event descriptions presented at the consensus meeting. A true estimate of inter-rater reliability would have involved two raters...
separately interviewing the same subject, rating the events, and presenting them to a panel for consensus. This remains an issue that needs to be examined in the future. Fourth, only half of the sample completed the LEDS interview, resulting in wider confidence intervals of the concurrent validity estimates. Ideally, all subjects would have been interviewed with the LEDS. However, due to the expense involved, this was not feasible. Fifth, the design of the study most likely negatively influenced our estimate of the test–retest reliability of the instrument. While the interviewer stressed the importance of answering all questions even though they had been discussed at the initial assessment period, youngsters were often distracted when participating in the phone interview and could be heard doing other things in the background (e.g. typing on a computer keyboard). It would have been preferable to administer the retest interview in person to circumvent these problems. Finally, having only 30 subjects as controls and 30 clinic cases with a range of psychiatric disorders limited our ability to examine any disease-specific hypotheses. Moreover, it would have been preferable to have a much larger sample size in both groups to definitively examine the psychometric properties of the SLES.

4.2. Inter-rater consensus reliability

In the current study, inter-rater consensus reliability was estimated by recording each rater’s ratings for objective threat, dependence and focus during the consensus meeting and comparing them with each other and with the final consensus ratings. Inter-rater consensus reliability was found to be moderate to almost perfect for ratings of objective threat ($\kappa$ ranged from 0.58 to 0.89) and substantial to almost perfect for dependence ($\kappa$ ranged from 0.80 to 0.96) and focus ($\kappa$ ranged from 0.91 to 0.97). The estimate of inter-rater consensus reliability for objective threat compares well with similar estimates of $\kappa$ (0.70–0.86) found using the LEDS among adults (Brown et al., 1987; Sherrill et al., 1997) and the PACE with children ($\kappa \geq 0.88$). Inter-rater consensus reliability for dependence among adults has also been examined: Paykel (1983) found 60% complete agreement for rating independence among raters, while Kendler and colleagues reported inter-rater agreement of $\kappa=0.79$ (Kendler et al., 1999). Another study of adults found inter-rater consensus reliability for focus ratings to be $\kappa=0.78$ (Parry et al., 1981). The results from this study indicate that the SLES has good inter-rater consensus reliability.

4.3. Test–retest reliability

The test–retest reliability of the instrument was also examined. Approximately 1 week after the initial interview, subjects were re-interviewed (over the telephone) about the occurrence of events during the 1-year period prior to the initial interview. Events occurring after the initial interview period were not included. Global analyses showed almost perfect retest reliability for total number of events (ICC=0.93) and substantial reliability for severe events (ICC=0.70), with adolescents (ICC=0.86) being more consistent in their reports of severe events compared with children (ICC=0.43). Our estimates of overall test–retest reliability are comparable to the test–retest reliability or $r=0.97$ reported in another study of young adolescents (Dise-Lewis, 1988) and $r=0.65$ reported in a sample of children (Lindsay, 1994). However, closer examination of the reporting of specific events gives a different picture. For all events, the test–retest reliability was found to be $\kappa=0.68$, which was significantly higher for females ($\kappa=0.74$ vs. 0.62) and adolescents ($\kappa=0.72$ vs. 0.61). The test–retest reliability was higher for severe events ($\kappa=0.80$) and was not significantly different for females or for adolescents. Few investigations have been reported that have examined the two levels of test–retest reliability. One of the few reports examining the overlap of specific events similarly found lower test–retest reliability estimates (median $r=0.25$) (Dise-Lewis, 1988). Glen et al. (1993), using the semi-structured Psychosocial Assessment of Childhood Experiences (PACE) interview, reported overall estimates of test–retest reliability to range from 0.63 to 0.76. Examining the overlap in
specific events reported at the initial and retest interviews, Glen et al. (1993) reported an identical overlap of approximately 50% which was comparable to the 54% overlap found for all events. However, in the current study, there was a significantly greater overlap in reporting of all events at initial and retest interviews for adolescents (59%) compared with children (47%). Although their sample included children aged 8 to 16 years, the small number of subjects included (n=15) precluded them from examining children’s and adolescents’ reports separately. For respondent-based instruments, the LEC has been reported to have good test–retest reliability (ICCs ~ 0.70); however, whether the same events were reported on at the two time periods was not examined (Brand and Johnson, 1982).

4.4. Discriminant validity

In this report, several aspects of the psychometric properties of the SLES were examined. As described above, the sample of 60 youngsters was evenly split into 30 with psychopathology and 30 without lifetime psychopathology. One objective of recruiting half of the sample with psychopathology was to inflate the base rate of ascertainable events to determine if the SLES was able to assess them. The other objective was to determine if the SLES-ascertained events discriminated between youngsters with and without psychopathology as has been shown in the literature (e.g. Goodyer, 1990; Williamson et al., 1995a). Children with psychopathology had significantly more total events, severe events rated at 3—moderate or 4—high objective threat, total objective stress, and total subjective stress. In addition, children with psychopathology were found to have significantly more behavior-dependent and behavior-independent events across all measures of stressful life events than normal controls. An increase in behavior-dependent events among children with psychopathology has been previously reported in at least two other reports (Sandberg et al., 1998; Williamson et al., 1995). Thus, the results of this study suggest that the SLES has adequate discriminant validity.

4.5. Concurrent validity

To determine how well the SLES compared with other instruments that assess environmental stressors (i.e. concurrent validity), the LEC and the LEDS were also administered. With regard to the LEC, initial comparisons showed that the SLES and the LEC had good agreement for both total events (ICC = 0.80) and severe events (ICC = 0.83). Indirectly, this finding converges with an earlier report showing that the LEC and the LEDS similarly discriminated between depressed and normal control adolescents (Duggal et al., 2000). As discussed by Gorman and colleagues (1993) and Katschnig (1986), the limitation in this approach for comparing events reported by different instruments is that it is possible that the two instruments, while reporting on the same number of events, are actually reporting on different events. Indeed, analyses comparing specific events showed much lower levels of agreement for both total events (κ = 0.28) and for severe events (κ = 0.38). With regard to the agreement for severe events, Duggal et al. (2000) similarly found that approximately 30% of the acute and chronic stressors were identified by both instruments (Duggal et al., 2000). It is not surprising that the SLES and the LEC have only slight agreement on the occurrence of specific events, and there are several reasons why the two do not concur. First, in the development of the SLES, potentially negative events were extracted from several life events instruments of which one was the LEC. As a result, the SLES contains over 30 more events than does the LEC. Second, for events occurring to others, the SLES probes for events occurring to close others and family members, while the LEC leaves open to interpretation the definition of a close family member or friend. Finally, the SLES probes the youngster to date as precisely as possible the occurrence of an event and uses memory cues (such as birth date and holidays) to pinpoint when an event has occurred, while the LEC does not. In the absence of memory cues, it is likely that the youngster indicated events occurring during the year prior to interview that actually happened outside of that year period.

Similarly, events ascertained with the SLES were compared with those ascertained from the
LEDS. Due to the amount of time and effort involved in administering the LEDS, only half of the sample was interviewed with the LEDS. As previously discussed (Williamson et al., 1998), the LEDS assesses a range of life events and 90% of the events assessed have little or no threat or only some threat. Further studies of adults (Brown and Harris, 1978, 1989), as well as adolescents (Williamson et al., 1998), have shown that high threat or so-called severe events are etiologically related to the onset of depression. Analyses comparing the SLES and the LEDS showed that the level of agreement was substantial between the two instruments for severe events ($\kappa = 0.77$). In fact, the source of disagreement between the two instruments was due entirely to more severe events being rated by the SLES but not the LEDS. All severe events rated by the LEDS were similarly ascertained by the SLES. We are unaware of any other report that has attempted to examine the concurrent or convergent validity of an interview-based assessment with the LEDS as only studies comparing the LEDS with respondent-based instruments have been reported (Katschnig, 1986; Gorman, 1993). The results of this study indicate that the SLES has substantial concurrent validity with the LEDS.

4.6. Comparison of child–parent reports

Finally, a parent for each child was interviewed about life events occurring to the child, and the parent’s report was compared with the child’s report. Global examination of child–parent reports showed substantial to almost perfect agreement for all events ($\text{ICC} = 0.81$) and severe events ($\text{ICC} = 0.75$). Child–parent agreement for severe events was found to be significantly higher for adolescents ($\text{ICC} = 0.80$ vs. 0.69). Closer examination of specific events found only moderate agreement for total events ($\kappa = 0.48$) and substantial agreement for severe events ($\kappa = 0.73$). Thus, for severe events only, the agreement between child–parent reports appears to be adequate to suggest that either informant can accurately report on events occurring to the youngster. Few reports in the literature have examined the issue of child–parent reports of stressful life events. In fact, most studies have not directly considered the issue of informant variability, relying instead on either the parent or the child alone to serve as the informant. Our findings are consistent with those reported by Monck and Dobbs (1985), who found that 93% of all severe events were co-reported by adolescent females and their mothers. One study examining a large sample of first-grade children reported a correlation of just 0.21 between parent and child reports after controlling for total number of events reported (Rende and Plomin, 1991). In addition, only moderate correlations between parent-child reports for daily stressors have been reported (Beasley and Kearney, 1996). One report of children found child–parent agreement to be approximately 70% (Lindsay, 1994), although events were considered globally and not examined individually. We suspect that using a systematic interview approach with specific probes such as the SLES can be expected to yield higher parent-child agreement, especially for significantly high threat stressors. Milder stressful experiences do not appear to be reliably reported by children or their parents. Investigators interested in examining the importance of milder stressors might want to consider assessing both parents and children.

Anecdotally, when interviewing parents about stressors in their child’s life, it is often necessary to remind the parent that you are asking about events that are focused on the child. Parents often want to talk about events that are stressors for themselves but that their son or daughter might not be directly aware of or involved in (e.g. marital problems, housing problems). While the current study indicates that there is good agreement between child–parent reports and especially for severe events among adolescents, it remains to be seen whether reports from different informants differentially relate to the onset and course of disorders.

4.7. Conclusion

Research examining the contribution of stressful environments to child and adolescent psychiatric disorders has been hindered by the available assessment methods. Due to convenience, self-report checklists have been primarily used, but
findings from these studies are somewhat limited in that stressors are broadly assessed, although depending on the research question, this might be sufficient (Duggal et al., 2000). Comprehensive investigator-based stressful life events assessments such as the LEDS represent a significant improvement in the assessment of environmental stressors, but their use is limited because they are time and labor intensive. In an effort to find a common ground between the two divergent methods for assessing stressful environments in children and adolescents, a new instrument was developed, the Stressful Life Events Schedule (SLES). In the current study, the psychometric properties of the SLES, specifically the discriminant validity, test–retest and concurrent validity of the SLES (particularly with the LEDS), have been reported. Research examining the effects of environmental stressors on the onset and course of illness can greatly benefit from a more comprehensive assessment of the environment such as that provided by the SLES.

The most recent version of the SLES and related material can be requested from the author.

Acknowledgments

This research was supported by grants P30 MH55123 (to Dr Brent), P01 MH41712 (to Dr Ryan), and K01 MH001957 (to Dr Williamson) from the National Institute of Mental Health.

References


