The Community Child Health Network Life Stress Interview: a brief chronic stress measure for community health research

Lynlee R. Tanner Stapleton, Christine Dunkel Schetter, Larissa N. Dooley, Christine M. Guardino, Jan Huynh, Cynthia Paek, Elizabeth Clark-Kauffman, Peter Schafer, Richard Woolard, Robin Gaines Lanzi & Community Child Health Network

To cite this article: Lynlee R. Tanner Stapleton, Christine Dunkel Schetter, Larissa N. Dooley, Christine M. Guardino, Jan Huynh, Cynthia Paek, Elizabeth Clark-Kauffman, Peter Schafer, Richard Woolard, Robin Gaines Lanzi & Community Child Health Network (2016) The Community Child Health Network Life Stress Interview: a brief chronic stress measure for community health research, Anxiety, Stress, & Coping, 29:4, 352-366, DOI: 10.1080/10615806.2015.1058368

To link to this article: https://doi.org/10.1080/10615806.2015.1058368
The Community Child Health Network Life Stress Interview: a brief chronic stress measure for community health research

Lynlee R. Tanner Stapleton, Christine Dunkel Schetter, Larissa N. Dooley, Christine M. Guardino, Jan Huynh, Cynthia Paek, Elizabeth Clark-Kauffman, Peter Schafer, Richard Woolard, Robin Gaines Lanz and Community Child Health Network

Department of Psychology, University of California, 1285A Franz Hall, Los Angeles, CA 90095-1563, USA; NorthShore University HealthSystem, 2650 Ridge Ave, Evanston, IL 60201, USA; New York Academy of Science, 1216 5th Avenue, New York, NY 10029, USA; Baby Love Plus Consortium, 1929 Mail Service Center Raleigh, NC 27699-1929, USA; School of Public Health, University of Alabama at Birmingham, 1665 University Blvd., 227 RPHB, Birmingham, AL 35294, USA; Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, 6100 Executive Blvd, Bethesda, MD 20892, USA

ABSTRACT

Background and Objectives: Chronic stress is implicated in many theories as a contributor to a wide range of physical and mental health problems. The current study describes the development of a chronic stress measure that was based on the UCLA Life Stress Interview (LSI) and adapted in collaboration with community partners for use in a large community health study of low-income, ethnically diverse parents of infants in the USA (Community Child Health Network [CCHN]). We describe the instrument, its purpose and adaptations, implementation, and results of a reliability study in a subsample of the larger study cohort. Design and Methods: Interviews with 272 mothers were included in the present study. Chronic stress was assessed using the CCHN LSI, an instrument designed for administration by trained community interviewers to assess four domains of chronic stress, each rated by interviewers. Results: Significant correlations ranging from small to moderate in size between chronic stress scores on this measure, other measures of stress, biomarkers of allostatic load, and mental health provide initial evidence of construct and concurrent validity. Reliability data for interviewer ratings are also provided. Conclusions: This relatively brief interview (15 minutes) is available for use and may be a valuable tool for researchers seeking to measure chronic stress reliably and validly in future studies with time constraints.

Background and objectives

Psychosocial stress is a major contributor to health (Contrada & Baum, 2011). Defined as demands that tax or exceed the resources of the organism (Cohen, Kessler, & Gordon, 1997), stressors can be distinguished as episodic or chronic. Episodic or acute stressors are relatively discrete events such as losing one's job, the sudden death of a loved one, or a tornado. In contrast, chronic stressors refer to long-term and ongoing difficulties such as living in an unsafe neighborhood, poverty, an unstable relationship with one's intimate partner, or exposure to racial discrimination. Although these two types of stress sometimes co-occur, chronic stressors appear to have particularly...
detrimental effects on health (McEwen, 1998; McGonagle & Kessler, 1990). Chronic stress contributes to poorer mental health, especially major depressive disorder (Bruce & Hoff, 1994; Hammen, Kim, Eberhart, & Brennan, 2009), and has adverse effects on physical health such as immune system dysfunction, cardiovascular disease, and mortality through accumulation of wear and tear on major body systems (McEwen & Seeman, 1999). Chronic stress is also thought to play a key role in health disparities due to race, ethnicity, or socioeconomic status (Dunkel Schetter et al., 2013), including birth outcomes and prenatal depression (Borders, Grobman, Amsden, & Holl, 2007; Misra, Strobino, & Trabert, 2010; Séguin, Potvin, St.-Denis, & Loiselle, 1995).

Chronic stress is of particular interest to researchers working in underserved and high-risk communities because chronic stressors are more common among the poor. Some recognized sources of chronic stress are living in communities with crime, crowding, noise, and air pollution; financial strain and food insecurity; lack of adequate or good quality health care; interpersonal relationship conflict and role strain; and racial discrimination (Lepore, 1997; Steptoe & Feldman, 2001; Taylor, Repetti, & Seeman, 1997; Williams, Neighbors, & Jackson, 2003). For those living in poverty, chronic stressors tend to co-occur, accumulate, and persist (Dunkel Schetter & Dolbier, 2011).

Given the current emphasis on chronic stress in health, insufficient research has been done in large-scale, community, or population studies. This may be partially due to reliance on self-report life event checklist measures that have well-known limitations (Hammen et al., 2009; Lepore, 1997). Most importantly, life event checklists fail to distinguish between episodic and chronic stressors, thus precluding investigations of the potentially unique effects of chronic stress exposure. Additionally, most life event checklists and severity ratings of events (e.g. how undesirable was an event) are subject to various interpretations by participants, and may be biased by the current emotional or cognitive state of the participant (Hammen, 2005). Finally, checklists fail to take into account the personal circumstances or context in which a given stressor occurs. For example, research has established that experiencing a natural disaster may have a very different impact depending on the context of the event such as amount of property damage, and damage or injury to family or friends (Nolen-Hoeksema & Morrow, 1991).

To address these issues, interview methods of stress assessment were developed quite some time ago (Brown, 1989), and research comparing interview and checklist methods has generally concluded that interviews are less subject to emotional and cognitive bias, more accurate in the recall and dating of stressors, and better predictors of mental health outcomes (Hammen, 2005). The pioneering interview method was the Life Events and Difficulty Schedule (LEDS; Brown, 1989; Wethington, Brown, & Kessler, 1995). The LEDS uses a semi-structured interview approach to elicit detailed information about various life events and chronic stressors, including their circumstances, time frame, and relation to other reported stressors. A panel of independent, trained raters then uses this information to rate each stressor, using a standard “dictionary”. This rating reflects the “objective” threat of a stressor – how an average person under identical circumstances would experience the event – but does not reflect how the participant emotionally reacted to the event. Thus, the chief advantage of the LEDS is that it assesses episodic and chronic stressors distinctly, with less bias, and obtains a rating of the impact of each stressor by within an individual’s personal context.

Guided by the LEDS, the UCLA Life Stress Interview (LSI; Hammen et al., 1987; Hammen, Marks, Mayol, & de Mayo, 1985) was developed in the 1980s as a relatively streamlined and accessible tool for the systematic assessment of episodic and chronic stressors. A key addition was the assessment of chronic, stressful conditions in several specified life domains (e.g. intimate partner relationships, work, and finances). The level of stress in each domain is rated by a trained interviewer using a five-point scale consisting of behaviorally specific anchor points. The UCLA LSI has been used widely with adult, child, and adolescent samples from community and clinical populations (Hammen, Ellicott, Gitlin, & Jamison, 1989; Hammen, Henry, & Daley, 2000; Hazel, Hammen, Brennan, & Najman, 2008), with adequate inter-rater reliability (Brand, Schechter, Hammen, Le Brocque, & Brennan, 2011; Hammen et al., 2009) and strong predictive validity. For example, chronic stress LSI scores in
early adulthood predicted later depression and physical health in past research (Hammen et al., 2009; Raposa, Hammen, Brennan, O’Callaghan, & Najman, 2014).

Despite the advances offered by better assessment methods such as the LEDS and the UCLA LSI, these methods have not been implemented widely in large-scale community or population health investigations for various reasons. Administration can be very time-consuming, and the semi-structured format requires a high level of training and skill with both general interviewing techniques and the instrument itself. Scoring procedures are laborious, involving thorough post-interview review, rating dictionaries, and team consensus. Together, these aspects can lead to substantial burden on respondents, interviewers, investigators, and their staff, which in turn eliminates feasibility for most large-scale research protocol, especially when stress assessment is one of several topics of the research. However, this method of assessment has potentially high value in such studies.

The current study chronicles our efforts to create a targeted, practical, and efficient method of assessing chronic stress in a large community research project on health disparities. For our study purposes, we shortened and adapted the chronic stress portion of the UCLA LSI for use by community interviewers administering it in the homes of new parents. In this paper, we describe the development of the revised LSI chronic stress interview instrument, developed as part of the Community Child Health Network (CCHN), demonstrate its feasibility for use in field and uncontrolled settings, and provide the first evidence of its reliability and validity in a subsample from the larger study.

CCHN is a five-site interdisciplinary research network formed by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) (see Acknowledgements for complete list of members). The goal of the network is to examine how community-, family-, and individual-level stressors and resources influence and interact with biological factors to affect maternal and child health in a diverse sample across five study sites and by actively using community partnership methods (Ramey et al., 2015). Community partnership methods entail collaboration between scientists and community partners in all stages of research (Israel, Eng, Schulz, & Parker, 2005), and have the advantages of improving participant recruitment and retention, rapport with participants, and cultural sensitivity (Minkler, 2005), as well as providing jobs and skill-building opportunities for community members.

To design the new CCHN LSI interview, we selected a few key domains that were prioritized by community partners but for which no suitable measures existed. We also developed a new domain of co-parenting to capture a relevant area of potential stress for parents of infants, a portion of which were unmarried or did not live together. Community partners and scientists devised standardized and simplified interview language appropriate to our community sample together with new scoring procedures for trained community interviewers. Once the study was completed, we selected a subsample of the CCHN cohort and conducted reliability tests on that group by having trained raters independently code either taped interviews or written transcripts of tapes. We also tested validity by using the larger CCHN dataset, which contained several other standardized measures of stress, mental health, and biomarkers.

Method

Design

CCHN conducted a longitudinal, observational study of mothers and a subset of fathers, described in detail elsewhere (Dunkel Schetter et al., 2013; Ramey et al., 2015). In brief, mothers of African-American, Latino and White race/ethnicity were recruited and provided written informed consent following childbirth in designated catchment areas in each of the five sites (with the exception of one site which recruited mothers prenatally). Recruitment efforts involved oversampling based on poverty status and risk for preterm delivery in order to identify an at-risk population. Fathers were invited to participate and provided separate informed consent if mothers agreed verbally for the study staff to contact them; only maternal data are used for the current study though fathers also
completed the CCHN LSI. Interviews were conducted in English or Spanish in homes, with attempts to match interviewer ethnicity to that of the participant. Initial assessments were conducted at approximately one month post-birth (T1) and about 6 months post-birth (T2) when the CCHN LSI was administered, with subsequent follow-up interviews up to 24 months of child’s age (data not included). Study procedures and protocols were reviewed and approved by the Institutional Review Boards of all community and academic institutions within CCHN.

Sample
A total of 1656 mothers (68% of the 2448 mothers enrolled at T1 in the CCHN study) completed assessments at 6 months after birth (T2). Demographically, this group did not differ from the full cohort. For the current inter-rater reliability analyses, a subsample of 272 mothers was selected systematically from this T2 dataset. Given variation in sample demographics across site, we randomly sampled a proportion of each ethnic group in each site. In a few cases (<5%), substitute cases were randomly chosen due to poor audio quality of the interview recording. The final subsample for reliability testing consisted of 100 African-American/Black mothers (37%), 100 non-Hispanic White mothers (37%), and 72 Hispanic/Latina mothers (26%). Table 1 provides demographic characteristics for the T2 CCHN sample of mothers and the reliability subsample. The subsample was largely representative of the T2 sample, with a few small magnitude differences; mothers in the reliability sample were slightly older, had more education, and had slightly higher household income.

Measures
Adapted chronic LSI (CCHN LSI)
The UCLA LSI contains both episodic and chronic stress portions; however, the chronic stress portion of the UCLA LSI (Hammen et al., 1987) was chosen as the basis for adaptation in consultation with the author and to suit CCHN study purposes and needs. Modifications to the original instrument were made in consultation with Hammen (personal communication, November 26, 2013, CDS and LTS) with the aims of reducing length and participant burden, increasing accessibility for respondents of diverse educational and cultural backgrounds, and practicality for administration and scoring by community interviewers in the field. Adaptations are described in detail here. The complete CCHN LSI instrument is available in online supplemental materials and additional information including interviewer training materials may be obtained from corresponding author.

Table 1. Demographic data for T2 sample of mothers and reliability subsample.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Full T2 sample (n = 1656)</th>
<th>Reliability sample (n = 272)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>25.7 (5.67)</td>
<td>26.8 (6.53)</td>
</tr>
<tr>
<td>M (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>18–42</td>
<td>18–42</td>
</tr>
<tr>
<td>Education (years)</td>
<td>13.0 (2.93)</td>
<td>14.0 (2.89)</td>
</tr>
<tr>
<td>M (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>2–26</td>
<td>7–23</td>
</tr>
<tr>
<td>Annual household income (per capita)</td>
<td>$14,781</td>
<td>$20,783</td>
</tr>
<tr>
<td>M</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>$26,959</td>
<td>$38,098</td>
</tr>
<tr>
<td>Relationship status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>32%</td>
<td>43%</td>
</tr>
<tr>
<td>Partnered</td>
<td>48%</td>
<td>36%</td>
</tr>
<tr>
<td>Single</td>
<td>20%</td>
<td>21%</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>54%</td>
<td>37%</td>
</tr>
<tr>
<td>Hispanic/Latina</td>
<td>23%</td>
<td>26%</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>23%</td>
<td>37%</td>
</tr>
</tbody>
</table>
**Content.** The CCHN LSI includes three of the eight original chronic stress domains: **neighborhood environment**, **family relationships**, and **partner relationships**. CCHN elected to omit the other domains either because they were less relevant to the study goals and population (e.g. social life) or were covered by other standardized measures in the protocol (e.g. finances). The team also developed a new domain of **co-parenting** stress for inclusion based on the importance of this experience for parents of infants. The **partner** and **co-parenting** sections included two parallel versions depending on whether the mother was currently in an intimate relationship or not, and if she was co-parenting with the baby’s father or a new partner. Of note, three of the four domains included in the CCHN LSI captured various interpersonal stressors that have been found to be most common and most upsetting in prior research (Paykel, 2003). Overall, these modifications made the CCHN LSI significantly briefer than the original UCLA LSI, requiring an average of 12 minutes to administer instead of 30 minutes.

The original UCLA LSI is loosely structured allowing for considerable interviewer discretion in wording, sequence of questions, and probing. For use in this context, we simplified and further standardized interviewer instructions, added questions, simplified terms, and further structured prompts and behavioral anchors to improve consistency and ease of administration and scoring. We also organized questions within each domain into subsections, and interviewers provided separate ratings on each subsection (see online supplement for complete instrument). Finally, we included modifications specific to our study population, such as referring to “relationship partner” rather than “spouse” and anchoring the time frame of inquiry around the birth of the most recent child (typically the past six months). Community partners and interviewers reviewed and provided feedback on these modifications during an initial pilot phase.

The interview was translated into Spanish for use with participants for whom Spanish was their primary language. This was done by a professional translating service experienced in survey research translation. The translation was then vetted by bilingual team members and pre-tested in the study sites before implementation. The Spanish translation of the CCHN LSI is available from the authors upon request.

**Training and administration.** The CCHN LSI was administered to all participants at the T2 visit. This portion of the protocol was audio-recorded with participant consent for later reliability and content analysis. Most CCHN interviewers were community members who did not have prior interviewing or clinical experience. They were therefore provided with training, including guidance and practice in interviewing techniques (e.g. addressing sensitive topics, responding objectively, and maintaining standard prompts), training in cultural competence, and guidance in fieldwork safety procedures.

Interviewers also completed specialized training in the CCHN LSI instrument. The first stage was conducted by the first author who trained project coordinators at each site in a centralized web conference and in-person meetings. Then individual interviewers watched recordings of these network meetings and received further site-specific training led by project coordinators, including rehearsal, role-plays with discussion, and establishment of scoring consistency. Questions or discrepancies were directed back to the first author for resolution. Ongoing site-specific trainings and meetings were also recommended to maintain reliability over time.

As in the original procedures, interviewers were allowed flexibility to probe to the degree needed to gather the necessary information, as well as to adapt their wording to the respondent’s needs, while covering all given sections. Interviewers were instructed to elicit information surrounding the factual circumstances in each domain and to disregard participant evaluations of or responses to the stressors, so as to maintain the objectivity of stress ratings.

**Scoring.** The UCLA LSI recommends that interviewers write summary paragraphs after the interview to assist in scoring, and interviewers are extensively trained in clinical interviewing for mental health research (Hammen et al., 1987). To assist our community interviewers with reliable scoring, we used a
two-step scoring process. After probing each domain subsection, interviewers provided a rating for that subsection according to specific behavioral anchors. All ratings were made on a five-point Likert-type scale (1 = exceptionally positive conditions, 5 = exceptionally negative conditions) by marking the appropriate checkbox, with half-step ratings (e.g. 3.5) permissible. An example for rating partner conflict specifies a rating of 1 as “rare or only minor conflicts with very good resolution”, and a rating of 5 as “frequent, intense conflicts that do not resolve and/or are abusive”. (See online supplemental material for further detail.) After completing all prompts and subsections, interviewers provided an overall rating for each domain, taking into account all given information (consistent with UCLA LSI procedures). Overall domain ratings were a summary judgment of chronic stress in that area of life, made by the interviewer in light of all the information they had obtained within that section, but they were not mathematical averages of the subsection ratings.

Of note, subsection ratings were not added to the protocol until after initial field interviews revealed the need (see Table 2 for subsection topics). Therefore, approximately 20% of the reliability subsample received only overall domain ratings. Adding subsection ratings did not increase the amount of time necessary to administer the interview; in fact, it seemed to reduce it. Also, for the brief co-parenting domain, only a single overall rating was made, as early testing determined that it was not possible to obtain sufficient information to make reliable subsection ratings within given time constraints.

Consistent with UCLA LSI procedures, the goal of ratings was to reflect objective assessments of ongoing stressful conditions without considering participants’ subjective reactions to the stressor, in order to minimize the potential impact of mood or cognitive biases. Throughout, interviewers could amend earlier scores if information provided later in the LSI interview revealed pertinent information.

**Additional measures used in validation**

Many additional standardized measures were administered to the full CCHN sample, some of which were used here to assess construct and concurrent validity of the CCHN LSI instrument. These are briefly described in the following (for more information, see Dunkel Schetter et al., 2013).

<table>
<thead>
<tr>
<th>Table 2. Means, standard deviations, and ICC coefficients for CCHN LSI field and laboratory ratings.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCHN LSI construct</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td><strong>Neighborhood domain</strong></td>
</tr>
<tr>
<td>Overall rating</td>
</tr>
<tr>
<td>Safety</td>
</tr>
<tr>
<td>Noise</td>
</tr>
<tr>
<td>Neighbors</td>
</tr>
<tr>
<td><strong>Family domain</strong></td>
</tr>
<tr>
<td>Overall rating</td>
</tr>
<tr>
<td>Availability</td>
</tr>
<tr>
<td>Closeness</td>
</tr>
<tr>
<td>Acceptance</td>
</tr>
<tr>
<td>Support</td>
</tr>
<tr>
<td>Conflict/resolution</td>
</tr>
<tr>
<td><strong>Partner domain – current partner</strong></td>
</tr>
<tr>
<td>Overall rating</td>
</tr>
<tr>
<td>Commitment</td>
</tr>
<tr>
<td>Closeness</td>
</tr>
<tr>
<td>Support/dependability</td>
</tr>
<tr>
<td>Conflict</td>
</tr>
<tr>
<td><strong>Co-parenting with baby’s father domain</strong></td>
</tr>
<tr>
<td>Overall rating</td>
</tr>
</tbody>
</table>

Note: ICC, intra-class correlations.
Stress questionnaires

Perceived stress. The Perceived Stress Scale (PSS; Cohen, Kamarck, & Meromelstein, 1983) measures the degree to which respondents perceive their lives over the past month as unpredictable, uncontrollable, and overburdened. The PSS was developed for use with community samples with at least an eighth-grade education. We administered the 10-item brief version (Cohen & Williamson, 1988) at T1. Items were answered on a 5-point scale (1 = never to 5 = almost always) based on how often the participant felt or thought a certain way. Responses were summed, with 4 positively worded items reverse-coded, for a total score from 10 to 50.

Life events. A life events checklist adapted from epidemiologic mental health research and used in several past maternal studies (Parker-Dominguez, Dunkel-Schetter, Mancuso, Rini, & Hobel, 2005) was administered at T1. Participants reported whether each of 24 events occurred in the past year, and how negative or positive the impact of each event was. For the current analyses, we scored the total number of life events only (Life Event Count) with a range of 0–24.

Prenatal stress. At T1, we included a standardized measure of overall prenatal stress as reported retrospectively. The original 12-item stress index of the Prenatal Psychosocial Profile (Curry, Campbell, & Christian, 1994) was revised into a shortened 10-item version (Misra, O’Campo, & Strobino, 2001). Items captured stress related to money worries, problems with family, work problems, and being generally overloaded during pregnancy and were answered on a 4-point rating scale (1 = no stress to 4 = severe stress). Scores were summed to produce a total prenatal stress score ranging from 10 to 40.

Interpersonal violence. A modified version of a standard screener for intimate partner violence was administered to mothers at T1. The HITS includes four items related to physical Hurt, Insult, Threats, and Screaming toward self or others in the household (Sherin, Sinacore, Li, Zitter, & Shakil, 1998), plus an additional item regarding domination or emotional control (O’Campo, Caughy, & Nettles, 2010). Mothers responded using a 5-point frequency format (1 = never to 5 = frequently), with responses summed for a total score from 5 to 25.

Mental and physical health

Depressive symptoms. Participants completed the Edinburgh Postnatal Depression Scale (EPDS) at T1 and T2 (Cox, Holden, & Sagovsky, 1987), a screening instrument validated for use during the first-year postpartum. Participants had the option of completing the EPDS verbally or via written questionnaire to enhance confidentiality and honest reporting. The EPDS queries 10 common depressive symptoms (e.g., feeling sad or miserable, not looking forward with enjoyment to activities, self-blame) experienced in the past 7 days. Items are rated on a 4-point scale ranging from 0 to 3, with higher ratings corresponding to greater symptoms, with total scores ranging 0–30. Only data from T2 is presented here.

Post-traumatic stress. The Post-traumatic Checklist–Civilian Version (PCL-C; Keane, Silberbogen, & Weierich, 2008) was used to assess symptoms of Post-traumatic Stress Disorder (PTSD) at T2. The PCL-C asks about symptoms in relation to “stressful experiences” and is intended for use with populations who may have symptoms due to multiple traumatic events. Items ask how often in the past month the respondent was bothered by each of the 17 core symptoms of PTSD including intrusive re-experiencing (e.g., “repeated, disturbing memories, thoughts, or images”), avoidance and numbing (e.g., “avoiding activities or situations because they remind you of a stressful experience”), and symptoms of hyperarousal (e.g., “feeling jumpy or easily startled”). Response options were given on a 5-point scale (1 = not at all to 5 = extremely), resulting in total scores from 17 to 85.

Generalized anxiety. Generalized anxiety, or excessive worry and non-specific anxiety, was assessed using a module from the Mini International Neuropsychiatric Interview (Sheehan et al., 1998) at T2. Using several guided questions (e.g., “Have you worried excessively or been anxious about several things over the past six months?”, “Are these worries present most days?”), this instrument determines whether a likely diagnosis of Generalized Anxiety Disorder (GAD) is present. Administration is discontinued once responses indicate GAD is not present. This resulted in a dichotomous variable indicating likely present or absent GAD diagnosis.
**Biological stress markers.** A composite measure of allostatic load which captures cumulative wear on the body’s systems resulting from stress (McEwen & Seeman, 1999) was calculated as an indicator of biological stress. The 10 biological markers included: body mass index, waist-to-hip ratio, systolic and diastolic blood pressure, pulse, high-sensitivity C-reactive protein (HS-CRP), hemoglobin A1c, high-density lipoprotein (HDL) cholesterol, total-HDL cholesterol ratio, and diurnal cortisol slope. The number of biomarkers with a value in the top quartile of the respective distribution was summed, resulting in a composite variable ranging from 0 to 10. For this study’s purposes, allostatic load was not calculated for participants with more than three missing biomarker data points (n = 163).

**Inter-rater reliability**

To assess the reliability of interviewer ratings in the field, undergraduate research assistants were extensively trained to transcribe and code audio recordings of the CCHN LSI field interviews. The first author conducted training for the research assistants over a period of approximately two months. Raters were assigned cases to code during training and discussed ratings during weekly or biweekly coding meetings in order to establish preliminary consensus. Group meetings were held until the group was able to agree on 80% or more of codes prior to discussion. All recordings in the study subsample were then randomly assigned to individual raters for coding independently. Half of the cases were coded by listening to audio recordings, while the other half were coded from typed transcripts. Raters were not permitted to code any interviews to which they had previously listened for transcription purposes.

Intra-class correlation (ICC) coefficients were calculated, comparing the ratings of field interviewers with those of lab-based research assistants using a one-way random effects model of absolute agreement (ICC(1): McGraw & Wong, 1996; Model 1: Shrout & Fleiss, 1979). The single-measure coefficient was used, as is appropriate when scores are not averaged across multiple raters. This is a conservative approach designed to account for the inability to correctly apportion intra-rater reliability error, given our methods of interviewer and lab-rater case assignment. Benchmark guidelines given by Landis and Koch (1977) were used to interpret coefficients (0.00–0.20 as *slight*, 0.21–0.40 as *fair*, 0.41–0.60 as *moderate*, 0.61–0.80 as *substantial*, and 0.81–1.00 as *almost perfect*). Given the limited number of participants without a partner relationship (n = 38) or who were co-parenting with a new partner (n = 8), it was not possible to obtain valid ICC estimates for these versions of the *partner* and *co-parenting* sections.

**Results**

**CCHN LSI descriptive data**

Means and standard deviations for domain and subsection ratings on the CCHN LSI are provided for the reliability subsample (n = 272) in Table 2. Overall, mean chronic stress ratings were in the low to moderate range but there was significant positive skew in the distributions. As reported elsewhere, poor, near-poor, and Black/African-American mothers in the larger study sample experienced greater chronic stress on the CCHN LSI when compared to non-poor and White/Caucasian or Latina/Hispanic mothers (Dunkel Schetter et al., 2013).

**Inter-rater reliability**

Reliability coefficients (ICC) were calculated separately for audio-based and transcript-based lab scores (Table 2). There was some variation in reliability, but all reliabilities were at least in the *moderate* range, and many in the *substantial* or *almost perfect* range. ICC coefficients ranged from .43 (family availability; transcript method) to .88 (partner support/dependability; audio method), and F tests indicated that all coefficients significantly differed from zero at p < .05. ICCs were typically higher for
ratings of specific aspects of domains than for overall summary ratings. For example, while the overall neighborhood reliability estimate (using audio recordings) was .65, the reliability of safety, noise, and familiarity/problems with neighbors subsection ratings were .79, .85, and .67, respectively. This suggests that stress ratings were most consistent when raters were asked to focus on a few key areas of assessment, rather than consider multiple areas at once. Regardless of domain, reliability estimates were also noticeably higher (average of 0.10 points) for codes made by listening to audio recordings compared to transcript-based codes, suggesting that the vocal tone and inflection of the interviewer and subjects in audio recordings contributes to improved reliability.

There was some variability in reliability across CCHN network sites despite extensive interviewer training and monitoring. Across sites, ICCs ranged from .46 to .78 for overall neighborhood ratings; from .44 to .78 for family ratings; from .51 to .76 for partner ratings; and from .48 to .93 for co-parenting ratings. Average ICCs across domain ratings within site ranged from .56 to .77. This variability could be a function of many between-site differences, including potential variations in frequency of interviewer reliability meetings, interviewer background, and/or respondent characteristics.

**Validity**

For CCHN larger research goals, a composite LSI score was computed averaging chronic stress across all four major domains (neighborhood, family, partner, and co-parenting). The correlations between the CCHN LSI and other indicators of stress were published within a broader description of the many stress indicators used by CCHN (Dunkel Schetter et al., 2013). For present purposes, this provides preliminary evidence of convergent validity. Specifically, in the full CCHN T2 sample (N = 1656), the composite LSI score was significantly correlated with perceived stress at the earlier T1 assessment one month after birth ($r = .28, CI = 0.23, 0.33$) and concurrently as measured at T2 six months after birth ($r = .37, CI = 0.33, 0.41$). The LSI composite was also associated with T1 reports of number of life events in the year preceding pregnancy ($r = .27, CI = 0.22, 0.32$), with prenatal stress ($r = .28, CI = 0.23, 0.33$), and reported interpersonal violence ($r = .28, CI = 0.22, 0.33$), all controlling for race/ethnicity and poverty. Thus, the CCHN LSI demonstrated consistent, moderate positive associations with several other well-validated measures of stress as we would expect.

To assess the CCHN LSI instrument’s ability to predict mental and physical health outcomes measured concurrently, correlations were calculated among all CCHN participants who provided relevant data at T2 between overall domain scores (as rated by field interviewers) and measures of depressive symptoms, post-traumatic stress symptoms, GAD diagnosis, and allostatic load. As seen in Table 3, all domain ratings on the LSI were significantly associated with all three indicators of mental health, with the exception of co-parenting with a new partner, which was

<table>
<thead>
<tr>
<th>Health indicator</th>
<th>Neighborhood (n = 1605)</th>
<th>Family (n = 1604)</th>
<th>Partner (n = 1305)</th>
<th>No partner (n = 269)</th>
<th>Co-parenting with a baby’s father (n = 1526)</th>
<th>Co-parenting with a new partner (n = 111)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPDS</td>
<td>.19**</td>
<td>.25**</td>
<td>.36**</td>
<td>.27**</td>
<td>.17**</td>
<td>.15</td>
</tr>
<tr>
<td>PCL-C</td>
<td>.22**</td>
<td>.31**</td>
<td>.36**</td>
<td>.26**</td>
<td>.20**</td>
<td>.15</td>
</tr>
<tr>
<td>GAD</td>
<td>.09*</td>
<td>.14**</td>
<td>.19**</td>
<td>.21**</td>
<td>.10**</td>
<td>.14</td>
</tr>
<tr>
<td>Allostatic load</td>
<td>.06*</td>
<td>.08**</td>
<td>.06*</td>
<td>.09</td>
<td>.08**</td>
<td>.12</td>
</tr>
</tbody>
</table>

Note: EPDS, Edinburgh Postnatal Depression Scale; PCL-C, Posttraumatic Checklist–Civilian; GAD, Generalized Anxiety Disorder diagnosis. ns vary due to the limited availability of biomarker data.

*p < .05.

**p < .01.
completed by only a small minority of participants. Correlation coefficients ranged from .09 to .36. CCHN LSI domain ratings were more strongly associated with depressive and post-traumatic stress symptoms than with GAD diagnosis. Correlations for the primary CCHN LSI domains (neighborhood, family, partner relationship, and co-parenting with baby’s father) with allostatic load, the biological indicator, were positive and statistically significant though smaller in magnitude (.06–.08). Allostatic load was not significantly associated with the two CCHN LSI domains of no partner and co-parenting with a new partner, which again applied only to a small portion of participants.

Conclusions

The current study describes the development and implementation of a chronic LSI adapted from established methods for use in our large-scale, community health study conducted in five regions of the USA. Preliminary evidence of its reliability and validity is promising. The process of instrument adaptation involved several innovations for the present purposes: (1) selecting fewer domains, mainly those on interpersonal chronic stress (neighborhood, family, partner relationship, and co-parenting), which substantially reduced the interview’s length; (2) adding more structure to the interview script; (3) revising language and terminology to be more accessible to a large range of participants; and (4) streamlining the scoring and documentation process for ease of use. These changes were conducted in consultation with the author of the UCLA LSI with feedback from the community interviewers during piloting. We also tailored the measure by selecting life domains that were most relevant to our study population; the domains we chose to investigate may be different than those that are most relevant in populations of different demographics. In the future, researchers may choose to include other relevant domains of the original UCLA LSI (e.g. close friendships, work, and health) given their time constraints and participant characteristics. Although additional refinement and validation of the instrument can be undertaken, the CCHN LSI, and the process by which it was developed and implemented, contributes to the literature on the measurement of chronic stress, especially in population and large-scale community research.

ICC estimates of reliability between field interviewers and lab-based research assistants ranged from moderate to strong, particularly when audio recordings were used to make ratings. Thus, when ratings about participants’ chronic life stress were made using similar procedures in the lab and in the field, community-based interviewers were able to make scoring decisions generally comparable to those of well-trained research assistants. The reliability estimates obtained in this study were comparable to field versus lab inter-rater reliability for the earlier UCLA LSI (estimates .55–.89; Brand et al., 2011; Hammen et al., 2009). This suggests that it is acceptable to use interviewers who are not necessarily clinically trained professionals for the present purposes, providing that they are trained thoroughly in the protocol and monitored during data collection. We estimate that approximately 10 hours of training were required, which is less in comparison to other stress interviews that require longer training for different more time-intensive ratings. For example, the LEDS historically required eight days of interviewer training (Wethington et al., 1995).

Of note, the method of coding had noticeable impact on inter-rater reliability. Specifically, the use of written transcripts for rating interviews generally resulted in poorer reliability estimates than listening to interview recordings. The reduced consistency between field and lab ratings suggests that paraverbal information, such as emphasis, tone of voice, and pauses in speech, are important to reliable scoring. As a result of this unexpected finding, we recommend that scoring occur in the moment by interviewers or using audio recordings of the interviews rather than relying on transcripts, as sometimes occurs in larger, centralized studies. This approach eliminates the need for costly and time-consuming transcription.

Another important finding was that subsection ratings within domains were generally more reliable than the overall domain ratings, which suggests that it is easier for interviewers to generate reliable ratings when interview information is broken up into meaningful segments rather than larger,
more diverse composite ratings. This is further supported by the fact that the co-parenting domain, which was the shortest section of the interview, was the domain with the best reliability. Future work to refine the interview and its scoring might build on this and improve the instrument further.

There was some variability in reliability across sites, perhaps due to variations in protocol practice, interviewer monitoring, or participant characteristics. The sites were individually responsible for verifying that interviewers were satisfactorily trained prior to conducting interviews in the field and for maintaining quality over time, which may have introduced site “drift” from original standards and resulted in more modest inter-rater agreement. Sites with interviewers who had pre-existing interviewing experience tended to produce more reliable ratings, which underlines the importance of general interviewing skills, though notably not necessarily clinical skills that were required or recommended for the original LSI. Interviewer qualifications for the CCHN LSI should include attention to detail, conscientiousness, and willingness to maintain protocol across interviews, and investigators should maintain regular monitoring and reliability checks over time (e.g. monthly meetings). Site-based differences in reliability may also reflect higher and more variable stress among poor and ethnic/racial minority participants (Dunkel Schetter et al., 2013). In particular, reliability was lowest in our rural eastern North Carolina site. Given the challenges associated with assessing diverse and high-stress populations, we recommend periodic re-training based on field experiences and discussion of unusual cases, to ensure consistent procedures. Cataloging and sharing this type of information would also be valuable across research teams.

Results include initial evidence for construct and concurrent validity. A strength of the CCHN LSI was that scores were positively and moderately correlated with several other validated and commonly used stress measures, including standardized scale assessments of perceived stress, life events, prenatal stress, and interpersonal violence. The fact that the CCHN LSI was only modestly correlated with these other stress measures indicates that, as intended, it is capturing overlapping but distinct components from these other stress measures. The CCHN LSI was also significantly correlated with scores on standardized measures of three aspects of mental health – depressive symptoms, post-traumatic stress symptoms, and GAD symptoms – and a composite index of biological stress (allostatic load) providing additional validity evidence. It is notable that the magnitude of the correlations of two LSI subscales with allostatic load was $r = .08$ ($p’s < .01$; Table 3), and though relatively small in magnitude, these are larger correlations than we found with other validated stress measures such as the PSS at two time points (T1 $r = .04$, $ns$ and T2 $r = .05$, $ns$), life events ($r = .05$, $ns$), pregnancy stress ($r = .03$, $ns$), and interpersonal violence ($r = .02$, $ns$). Thus, this new chronic stress measure may be useful in studies of health that require shorter assessments than the much longer UCLA LSI interview. A further strength was the development of the interview using community partnership methods which afforded better ecological validity and cultural sensitivity. Divergent validity of the instrument was not explicitly examined, although the low-to-moderate magnitude correlations with other stress measures and health variables suggest that the LSI captures both common and unique variance. Moreover, the magnitude of the correlations with other stress measures was generally higher ($r’s$ ranged from .28 to .37) than correlations with the mental health measures ($r’s$ ranged from .09 to .36), as would be expected.

Limitations of this study include slight demographic differences between our reliability subsample and the T2 sample. Despite this, a reasonable distribution of scores was obtained within the reliability subsample. Finite study resources also limited the size of this validation sample, although the marginal benefit of a larger sample size may be modest. Additional work is needed to clarify which sections of the instrument are most useful for predicting specific health outcomes and to refine instrument content to maximize its usability and utility, affirm reliability, and confirm divergent validity. Specifically, interview sections that were completed by only a minority of participants (e.g. no partner relationship, co-parenting with new partner) and the newly added co-parenting section require further attention. As a preliminary step toward validating the co-parenting with new partner domain, we combined scores for the limited number of participants with scores for this domain ($n=8$) with scores for the larger number of participants with scores for co-parenting
with the baby’s father \((n = 264)\). The reliability estimates obtained using the combined co-parenting scores (ICC = .84 for lab audio coding; ICC = .67 for transcription coding) were nearly equivalent to those reported above for co-parenting with baby’s father (Table 2).

In conclusion, chronic stress is a potentially powerful construct in our growing understanding of social determinants of disease, and it is essential to understanding health disparities related to income and ethnic minority status. Yet few validated tools exist for effectively and efficiently assessing chronic stress in community populations. Thus, the development of a standardized, culturally sensitive, time-efficient, and cost-effective measure of chronic stress is a step forward for community health research. The adaptation of the CCHN LSI demonstrates how established, more elaborate assessment tools such as the UCLA LSI can be tailored to specific populations and research goals without compromising psychometric qualities. Strong instrumentation is needed to ensure adequate power for testing complex biopsychosocial stress models, especially mediational models of mechanisms by which stress affects health, and to increase an understanding of health disparities. While the current study evaluates the CCHN LSI for use with parents, we are confident that this new instrument has potential for use with a broad range of populations and research questions. Future studies can test the instrument in different samples and languages, as well as further refine and standardize the methods and training to provide an even better quality assessment of chronic stress. Ultimately, this work has the potential to advance our understanding of how and for whom chronic stress impairs physical and mental health and what can be done about it.

**Acknowledgements**

We wish to sincerely thank the work of our many community interviewers, study participants, and research assistants who contributed to this work. CCHN reflects joint endeavors of five local sites:

**Baltimore: Baltimore City Healthy Start and Johns Hopkins University**
- Community PI: M. Vance
- Academic PI: C. S. Minkovitz
- Co-Is: P. O’Campo, P. Schafer
- Project Coordinators: N. Sankofa, K. Walton

**Lake County, Illinois: Lake County Health Department and Community Health Center and the NorthShore University Health System**
- Community PI: K. Wagenaar
- Academic PI: M. Shalowitz
- Co-Is: E. Adam, G. Duncan*, A. Schoua-Glusberg, C. McKinney, T. McDade, C. Simon
- Project Coordinator: B. Clark-Kaufman

**Los Angeles: Healthy African American Families, Cedars-Sinai Medical Center, University of California, Los Angeles**
- Community PI: L. Jones
- Academic PI: C. Hobel
- Co-Pls: C. Dunkel Schetter, M. C. Lu
- Co-I: B. Chung
- Project Coordinators: F. Jones, D. Serafin, D. Young

**North Carolina: East Carolina University, NC Division of Public Health, NC Eastern Baby Love Plus Consortium, and University of North Carolina, Chapel Hill**
- Community PIs: S. Evans, J. Ruffin, R. Woolard
- Academic PI: J. Thorp
- Co-Is J. DeClerque, C. Dolbier, C. Lorenz
- Project Coordinators: L. S. Sahadeo, K. Salisbury

**Washington, DC: Virginia Tech Carilion Research Institute, Virginia Tech, and Washington Hospital Center, and Developing Families Center**
- Community PI: L. Patchen
- Academic PI: S. L. Ramey
- Academic Co-Pl: R. Gaines Lanzi
- Co-Is: L. V. Klerman, M. Miodovnik, C. T. Ramey, L. Randolph
- Project Coordinator: N. Timraz
- Community Coordinator: R. German
Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The Community Child Health Network (CCHN) is a community-based participatory research network supported through cooperative agreements with the Eunice Kennedy Shriver National Institute of Child Health and Human Development [U HD44207, U HD44219, U HD44226, U HD44245, U HD44253, U HD54791, U HD54019, U HD44226-05S1, U HD44245-06S1, R03 HD59584] and the National Institute for Nursing Research [U NR008929].

Supplemental data

Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/10615806.2015.1058368.

References


