Evidence-Based Review of Subjective Pediatric Sleep Measures

Amy S. Lewandowski, PhD, Marisol Toliver-Sokol, MD, and Tonya M. Palermo, PhD

1Department of Anesthesiology and Pain Medicine, University of Washington School of Medicine,
and 2Department of Psychiatry, Oregon Health & Science University

All correspondence concerning this article should be addressed to Amy S. Lewandowski, PhD, Seattle Children’s Hospital Research Institute, P.O. Box 5371, Seattle, WA, 98145-5005, USA.
E-mail: amy.lewandowski@seattlechildrens.org

Received September 22, 2010; revisions received November 19, 2010; accepted December 9, 2010

Objective This manuscript provides an evidence-based psychometric review of parent and child-report pediatric sleep measures using criteria developed by the American Psychological Association (APA) Division 54 Evidence-Based Assessment (EBA) Task Force. Methods Twenty-one measures were reviewed: four measures of daytime sleepiness, four measures of sleep habits/hygiene, two measures assessing sleep-related attitudes/cognitions, five measures of sleep initiation/maintenance, and six multidimensional sleep measures. Results Six of the 21 measures met “well-established” evidence-based assessment criteria. An additional eight measures were rated as “approaching well-established” and seven were rated as “promising.” Conclusions Overall, the multidimensional sleep measures received the highest ratings. Strengths and weaknesses of the measures are described. Recommendations for future pediatric sleep assessment are presented including further validation of measures, use of multiple informants, and stability of sleep measures over time.

Key words adolescents; children; sleep; systematic review.

Introduction

Sleep problems are commonly reported among youth in the general population (Johnson, Roth, Schultz, & Breslau, 2006; Smaldone, Honig, & Byrne, 2007). Many of these are behavioral in nature such as problems with sleep initiation or maintenance, poor sleep quality, and poor sleep habits or hygiene (LeBourgeois, Giannotti, Cortesi, Wolfson, & Harsh, 2005; Meltzer & Mindell, 2006). Problems falling asleep and staying asleep are reported by 11–47% of youth (Liu & Zhou, 2002; Russo, Bruni, Lucidi, Ferri, & Violani, 2007). Disturbed sleep is associated with problems with cognitive functioning, learning, attention, and school performance (Sadeh, Gruber, & Raviv, 2002, 2003; Wolfson & Carskadon, 2003). Poor sleep has also been linked with socioemotional problems particularly symptoms of anxiety and depression, behavior problems, and substance abuse in youth (Johnson & Breslau, 2001; Smedje, Broman, & Hetta, 2001). Nighttime sleep disturbances can lead to daytime sequelae including sleepiness and reduced functioning (Fallone, Owens, & Deane, 2002). Researchers have also shown associations between disrupted sleep and increased somatic complaints and poor health-related quality of life (Hart, Palermo, & Rosen, 2005).

Accurate assessment of sleep disturbance and associated behaviors has practical applications in research, clinical care, and measuring responses to sleep interventions. Measurement of sleep is also important for describing differences in pediatric populations (e.g., medical or psychiatric populations). Due to its multidimensional nature, a variety of measurement tools are utilized in sleep assessment. Polysomnography records the biophysiological changes that occur during sleep (e.g., brain function, heart rate, eye movement, muscle activation) to aid in the diagnosis of sleep disorders. Polysomnography is particularly useful for examining sleep staging, respiration, and limb movements during sleep and is currently the gold standard for objective sleep assessment (Marcus, 2001). Sleep patterns are also assessed with actigraphy.
through an actimetry sensor that continually records motor movement. Actigraphy has been shown to be a reliable and valid assessment of sleep in pediatric populations (Acebo et al., 1999) and it provides data on sleep patterns. While useful, both polysomnography and actigraphy have limitations. Polysomnographic assessment is lab based and does not measure sleep habits in the natural setting. Furthermore, while polysomnography and actigraphy provide information on sleep and sleep patterns, they fail to identify behavioral sleep disturbances (e.g., bedtime resistance, insomnia) or reasons for nighttime movements (e.g., sleepwalking).

Parent and child-report sleep questionnaires are a critical component of behavioral and physiological sleep assessment. These questionnaires are primarily retrospective with youth or their parents reporting on typical sleep patterns, disturbances, or behaviors (e.g., sleep habits/hygiene, sleep quality) over a specified time period (e.g., 1 week, 1 month). Questionnaire measures can be used alone or in conjunction with other sleep assessment tools to provide a comprehensive examination of sleep in youth.

During the past 20 years, the number of pediatric sleep measures being used in pediatric psychology research has increased significantly. Prior to this, psychologists examining pediatric sleep primarily used assessment tools designed for use with adults that were later adapted for youth. While some adult measures have documented reliability in pediatric populations, the adaptations often do not capture the unique characteristics of pediatric sleep. Because sleep architecture, patterns, and behavior evolve significantly from infancy through childhood and adolescence (Kahn, Dan, Groszwater, Franco, & Sottiaux, 1996; Sadeh, Raviv, & Gruber, 2000; Yang, Kim, Patel, & Lee, 2005) developmentally appropriate assessment is critical. Moreover, in pediatric sleep, the social context in which sleep occurs is important and questionnaires appropriate for use by proxy reporters are needed for very young children.

Measurement categories

This review focuses on the psychometric properties of questionnaires used to assess sleep-wake patterns and behaviors in children and adolescents; particularly measures of daytime sleepiness, sleep habits and hygiene, attitudes and cognitions associated with sleep, sleep initiation and sleep maintenance, and multidimensional measures. Measures of daytime sleepiness describe perceptions of drowsiness during wake hours, periods of reduced alertness, and tendency to fall asleep during the day. Sleep habits and hygiene are behavioral dimensions of sleep that include: bedtime routines, sleep environment (e.g., whether the child sleeps alone, darkness of room), and activities prior to sleep initiation (e.g., sleep schedule, screen time in bed). Sleep initiation/maintenance describe ease or difficulty with sleep onset, sleep duration, nighttime wakeings, sleep satisfaction, and depth of sleep. Sleep-related attitudes and cognitions describe cognitions related to the sleep experience such as thoughts and arousals at bedtime. Finally, multidimensional sleep measures are broad tools that measure several sleep domains. These measures commonly screen children for both physiological (e.g., sleep disordered breathing) and behavioral (e.g., insomnia) sleep problems. A review of other sleep assessment tools (e.g., polysomnography, actigraphy) and questionnaires exclusively focused on physiological sleep problems is beyond the scope of the current review.

Current review

The current review summarizes the evidence base for 21 pediatric parent and child-report sleep measures to guide pediatric psychologists in their use. Although previous review articles (Lomeli et al., 2007; Meltzer & Mindell, 2006) have summarized commonly used sleep assessment tools, these reviews failed to review the psychometric properties or the evidence base. In addition, Lomeli et al.’s (2007) review was published in a Spanish language journal with limited circulation. The current review fills this gap by presenting a comprehensive examination of the content and psychometric properties of pediatric sleep questionnaires. Specifically, the goals were to: (a) review the psychometric characteristics of sleep questionnaires used in pediatric psychology literature; (b) categorize evidence for each measure using American Psychological Association (APA) Division 54 Evidence-Based Assessment (EBA) Task Force criteria (Cohen et al., 2008); (c) describe each measures’ utility for different populations; (d) offer perspectives on the measures’ strengths and weaknesses; and (e) provide recommendations for future research and measure development in this area.

Method

Measure selection

Electronic searches of the Cochrane Database of Systematic Reviews, Medline, PubMed, EMBASE, CINAHL, and PsychINFO were conducted by two authors (M.T.S., A.L.). The time period for the searches was from January 1990 through January 2010. Subject headings included “sleep,” “assessment,” “measurement,” “infant,” “child,” “adolescent,” “pediatric” as well as specific sleep terms “sleepiness,” “sleep habits,” “sleep hygiene,” “pre-sleep arousal,” and “sleep disturbance” and expanded versions of these
terms. Due to difficulty locating original papers published prior to 1990, we made the decision to use 1990 as the beginning date. Reference lists of retrieved papers were scanned for additional citations.

Measures eligible for review had to meet the following criteria: (a) published in a peer-reviewed journal from January 1990 through January 2010, (b) the questionnaire was developed to assess parent and child-reported sleep disturbances in children/adolescents, (c) measures initially developed for adults were only included if they were later adapted and validated with child or adolescent populations, and (d) the measure was published in English.

A total of 40 measures were identified using initial search criteria. Eight measures were excluded because they were originally developed as adult measures, and reliability or validity with pediatric sleep populations was not established. Two measures were not included because they were developed exclusively to assess physiological sleep disturbances (e.g., obstructive sleep apnea). Five pediatric sleep measures were excluded because reliability or validity information was not published in a peer-reviewed journal. Two measures were excluded because they were duplicates or subscales of full-length measures included in this review. Finally, a citation for one measure published prior to 1990 was located in the search, however, because it was never utilized after that date it was excluded from the review. The final inclusion list had 21 measures. Based on their content, measures were then rationally group into five categories: sleep hygiene and habits, sleep initiation and maintenance, daytime sleepiness, sleep-related cognitions, and multidimensional sleep.

**Framework and assessment criteria**

We used evidence-based criteria developed by the Society of Pediatric Psychology Assessment Task Force (Cohen et al., 2008) to critique and categorize pediatric assessment measures. The criteria are based on: (a) the existence of validity and reliability data of the measure, (b) availability of the measure with instructions on its use and in scoring, and (c) use of the measure by other investigators with findings published in a peer-reviewed journal. Based on these criteria, measures were rated as (a) well-established, (b) approaching well-established, or (c) promising, assessment tools. In addition to sound psychometric properties, in order to achieve a “well-established” rating, it was required that the sleep measure had been used by two or more investigators/ investigative teams with ready access to the measure for use by other investigators. An “approaching well-established” categorization required that the measure had been utilized by at least two investigators, had moderate or vague psychometric properties, and access to the measure is available. Lastly, a “promising assessment” categorization was made when the measure had been described in at least one other peer-reviewed article, moderate or vague psychometric properties were presented, and access to the measure is available.

**Methodology of review process**

After identifying the 21 measures included for review, articles were obtained that provided psychometric data for each measure. Specifically, three types of reliability (internal consistency, test–retest, and cross informant) and two types of validity (construct and criterion related) were used in the primary assessment of the measures. Information on sensitivity and specificity and results of factor analyses are also included where available. Data extraction forms were used to summarize each measure. Table I summarizes the evidence-based ratings and reliability and validity data obtained from primary validation studies. Two independent raters (A.L., M.T.S.) reviewed the data to judge each measure based on EBA criteria. Raters demonstrated agreement in the substantial range ($\kappa = .86$) on evidence based ratings for 19 of the 21 measures. In the case of the two measures with discrepant ratings, the senior author (T.M.P.) adjudicated the final classification.

**Review and description of measures**

The data from this review are summarized in Table I and includes: age ranges, response formats, psychometric properties, and EBA ratings for each measure. The measures are grouped by the five identified measurement categories and are listed alphabetically by measure within each category. Data presented in Table I are drawn from the primary manuscripts that present psychometrics data. For the minority of measures, particularly those developed for another population (e.g., adults), data from subsequent publications are included.

**Measures of sleep hygiene and sleep habits**

Four measures of sleep habits were evaluated: the Adolescent Sleep Hygiene Scale (ASHS) (LeBourgeois et al., 2005), the Family Inventory of Sleep Habits, (Malow et al., 2009) the Bedtime Routines Questionnaire (BRQ) (Henderson & Jordan, 2010), and the Children’s Sleep Hygiene Scale (CSHS) (Harsh, Easley, & LeBourgeois, 2002). This group of measures examines sleep hygiene and sleep habits such as bedtime routines, activities surrounding bedtime, and the sleep environment. The four measures are all relatively short (12–31 items), with the ASHS and the BRQ also including domain or subscale scores. In terms of psychometric properties, each measure has published validity information which...
Table 1. Psychometric properties of subjective sleep measures

<table>
<thead>
<tr>
<th>Measure domains</th>
<th>Age</th>
<th>Respondents and format</th>
<th>Psychometric reliability</th>
<th>Psychometric validity</th>
<th>EBA rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adolescent Sleep Wake Initiative (ASWS)</td>
<td>12-18 years</td>
<td>28 item; adolescent report; 1 month retrospective report</td>
<td>Total: $\alpha = .86$ subscales: $\alpha = .64-.82$</td>
<td>Associations with ASKS</td>
<td>Approaching well-established</td>
</tr>
<tr>
<td>Children's Sleep Wake Initiative (CSWS)</td>
<td>2-8 years</td>
<td>40 item; Parent report; 1 month retrospective report</td>
<td>Total: $\alpha = .89$ subscales: $\alpha = .71-.89$</td>
<td>Factor analysis</td>
<td>Promising</td>
</tr>
<tr>
<td>Infant Sleep Questionnaire (ISQ)</td>
<td>12-18 months</td>
<td>10 items; Parent report; 1 month retrospective report</td>
<td>Test-retest: .78 (Richman criteria); .93 (maternal criteria)</td>
<td>Good sensitivity and specificity for identifying sleep problems</td>
<td>Well-established</td>
</tr>
<tr>
<td>Sleep and Settle Questionnaire (SSQ)</td>
<td>6 weeks to 6 months</td>
<td>34 items Parent report 1 subscale; 1 week retrospective report</td>
<td>Test-retest: &quot;moderate&quot; .39-.80 (Bother composites) Test-retest: .14-.76 (individual items)</td>
<td>Differentiates between mothers who did/did not report sleep problems; Sensitive to change/treatment effects</td>
<td>Promising</td>
</tr>
<tr>
<td>Tayside Children's Sleep Questionnaire (revised subscale of SDSC)</td>
<td>1-3 years</td>
<td>10 items; Parent report; 3 month retrospective report</td>
<td>Total: $\alpha = .85$ item-total correlations: .30-.72</td>
<td>Factor analysis; criterion cut-score</td>
<td>Approaching well-established</td>
</tr>
<tr>
<td>Daytime Sleepiness Questionnaire (CASQ)</td>
<td>11-17 years</td>
<td>16 items; adolescent report; &quot;When you might feel sleepy during a usual week&quot;</td>
<td>Total: $\alpha = .89$</td>
<td>Correlations and convergent validity with PDSS and SSHS; detected group differences between normal and clinical populations</td>
<td>Promising</td>
</tr>
<tr>
<td>Epworth Sleepiness Scale—Revised for Children (ESS)</td>
<td>2-18 years</td>
<td>8 items; Child and parent report; Present timeframe &quot;how likely are you/your child?&quot;</td>
<td>Total: $\alpha = .75$</td>
<td>ESS scores of children with S-SDB higher than controls; ESS scores associated with self-report anxiety and general health (in adolescents)</td>
<td>Approaching well-established</td>
</tr>
<tr>
<td>Pediatric Daytime Sleepiness Scale (PDSS)</td>
<td>11-13 years</td>
<td>8 items; child-adolescent report; present timeframe &quot;how often do you?&quot;</td>
<td>Total: $\alpha = .80$ Split-half: $\alpha = .81$</td>
<td>Factor analysis; significant relationship with total sleep time; sleepiness predicted poor academic performance, illness</td>
<td>Well-established</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Name of measure</th>
<th>Measure domains</th>
<th>Age</th>
<th>Respondents and format</th>
<th>Psychometric reliability</th>
<th>Psychometric validity</th>
<th>EBA rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher’s Daytime Sleepiness Questionnaire (TDSQ)</td>
<td>Daytime behaviors associated with sleep disturbance observed in classroom (e.g., difficulty staying awake, yawning)</td>
<td>4–10 years</td>
<td>10 items Teacher report; 1 week retrospective survey</td>
<td>Total: $\alpha = .80$</td>
<td>Correlated with daytime sleepiness subscale of CSHQ; factor analysis; total scores higher for children with sleep problems</td>
<td>Approaching well-established</td>
</tr>
<tr>
<td>Sleep Habits/Hygiene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adolescent Sleep Hygiene Scale (ASHS) (LeBourgeois et al., 2005; modified from Children’s Sleep Hygiene Scale)</td>
<td>Sleep inhibiting and sleep facilitating practices</td>
<td>12–18 years</td>
<td>28 items; 9 domain scores; adolescent report; 1 month retrospective survey</td>
<td>Total: $\alpha = .80$ subscales: $\alpha = .46–.71$</td>
<td>Associations with ASWS</td>
<td>Approaching well-established</td>
</tr>
<tr>
<td>Bedtime Routines Questionnaire (BRQ) (Henderson &amp; Jordan, 2010)</td>
<td>Children’s weekday and weekend bedtime routines; reactivity to routine changes, activities before bedtime</td>
<td>2–8 years</td>
<td>31 items; 4 subscale scores; parent report; present timeframe “how often do you?”</td>
<td>Total: $\alpha = .88$ subscales: $\alpha = .69–.90$</td>
<td>Factor analysis; Subscale items correlated with CRQ, CSHS &amp; CSWS; differentiates between good and poor sleepers</td>
<td>Promising</td>
</tr>
<tr>
<td>Children’s Sleep Hygiene Scale (CSHS) (LeBourgeois &amp; Harsh, 2001; Harsh et al., 2002)</td>
<td>Activities surrounding sleep; bedtime routines; stable bedtime and wake time</td>
<td>2–8 years</td>
<td>17 items; parent report; 1 month retrospective survey</td>
<td>Total: $\alpha = .76$</td>
<td>Associations with BRQ and CRQ</td>
<td>Promising</td>
</tr>
<tr>
<td>Family Inventory of Sleep Habits (Malow et al., 2009)</td>
<td>Sleep hygiene in children with autism (daytime and pre-bedtime habits, bedtime routine, sleep environment)</td>
<td>3–10 years</td>
<td>12 items; parent report; 1 month retrospective report</td>
<td>Children with ASD: $\alpha = .61$; typically developing: $\alpha = .53$ test–retest: children with ASD: $\alpha = .82$, typically developing: $\alpha = .56$</td>
<td>Subscales showed negative correlations with CSHQ relationships with SDB, sleep onset delay, and parasomnias for children not taking psychotropic medication only</td>
<td>Promising</td>
</tr>
<tr>
<td>Cognitions/ Beliefs about Sleep</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysfunctional Beliefs about Sleep Questionnaire (DBAS) (Gregory et al., 2009; adapted from Morin et al., 1993)</td>
<td>Beliefs and attitudes about sleep</td>
<td>8–10 years</td>
<td>24 items; 5 subscales; child report; present timeframe</td>
<td>Total: $\alpha = .76$ subscales: $\alpha = .09–.63$</td>
<td>Total DBAS score and some subscales predicted child-report SSR total score and SSR insomnia</td>
<td>Promising</td>
</tr>
<tr>
<td>Presleep Arousal Survey for Children (PSAS-C) (Gregory et al., 2008, revised scale by Nicassio et al., 1985)</td>
<td>Presleep arousal, physiologic and mental process prior to sleep</td>
<td>8–10 years; later used with adolescents up to 14 years</td>
<td>16 items; 2 subscales (cognitive and somatic); child report; present timeframe</td>
<td>Total: $\alpha = .85$ subscales: $\alpha = .75$</td>
<td>Correlate with anxiety and sleep measures; differentiate between clinical and community samples</td>
<td>Approaching well-established</td>
</tr>
</tbody>
</table>

*Table I. Continued*
<table>
<thead>
<tr>
<th>Name of measure</th>
<th>Measure domains</th>
<th>Age</th>
<th>Respondents and format</th>
<th>Psychometric reliability</th>
<th>Psychometric validity</th>
<th>EBA rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multi-dimensional</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brief Infant Sleep Questionnaire (BISQ) (Sadeh, 2004)</td>
<td>Brief questionnaire for screening infant and toddler sleep problems (sleep duration, night waking)</td>
<td>0–29 months</td>
<td>13 items; parent –report; 2 week retrospective report</td>
<td>Test–retest: (.82–.95) for all individual items</td>
<td>Correlated with actigraphy and daily sleep logs; moderate correlation with caregiver report of sleep; sensitivity in documenting developmental sleep trends</td>
<td>Well-established</td>
</tr>
<tr>
<td>Children’s Sleep Habits Questionnaire (CSHQ) (Owens et al., 2000; Preschool/toddler version, Goodlin-Jones et al., 2008)</td>
<td>Screener for behaviorally and medically based sleep problems (e.g., bedtime behavior, nightwakings, parasomnias, sleep-disordered breathing)</td>
<td>4–10 years; additional validation sample: 2.5–5 years</td>
<td>35 items; 8 subscales; parent report; 1 week retrospective report</td>
<td>Total: $\alpha = .68$ (community); $\alpha = .78$ (clinical) subscales; $\alpha = .36–.70$ (community); $\alpha = .56–.93$ (clinical)</td>
<td>ROC = .41, distinguish between clinical and control groups; sensitivity = .80, specificity = .72</td>
<td>Well-established</td>
</tr>
<tr>
<td>Pediatric Sleep Questionnaire (PSQ) (Chervin et al., 1997; Chevrin et al., 2000)</td>
<td>Presence of sleep-related breathing disorders, daytime sleepiness, snoring, inattention</td>
<td>2–18 years</td>
<td>69 items; parent report; 8 subscales; “When sleeping does your child…”</td>
<td>Subscale $\alpha$’s = .77–.89</td>
<td>PSQ snoring score predicts SRBD; restless leg score predicted PSG abnormality; factor analysis, sensitivity and specificity data on SRBD subscale</td>
<td>Well-established</td>
</tr>
<tr>
<td>Sleep Disturbance Scale for Children (SDSC) (Bruni et al., 1996)</td>
<td>Sleep initiation and maintenance, daytime sleepiness, sleep disordered breathing, sleep arousal</td>
<td>5–15 years</td>
<td>26 items; 6 subscales; parent report; 6 month retrospective report</td>
<td>Total: $\alpha = .79$ (control); $\alpha = .71$ (sleep disorder); test–retest .71 (total), .21–.66 (single items)</td>
<td>Factor analysis; total and factor scores differentiated clinical and control groups; good diagnostic accuracy (AUC = .91); SRDB subscale validated with actigraphy (Portuguese version)</td>
<td>Well-established</td>
</tr>
<tr>
<td>Sleep Habits Survey (SHS) (Wolfson &amp; Caskadon, 1998, Wolfson et al., 2003) (Modified: Gionnotti et al., 2002)</td>
<td>Self-reported usual sleeping and waking on school and weekend nights; School performance, daytime sleepiness, Sleep/ wake behavior problems</td>
<td>10–19 years</td>
<td>63 items; 3 subscales; adolescent report; 2 week retrospective report</td>
<td>Sleepiness scale $\alpha = .70$; sleep/ wake behaviors scale $\alpha = .75$</td>
<td>Correlated with both diary and actigraphy variables; associations were greater for school-night variables than weekend nights</td>
<td>Approaching well-established</td>
</tr>
<tr>
<td>Sleep Self Report (SSR) (Owens et al., 2000)</td>
<td>Sleep habits; problems falling asleep, sleep duration, night waking, daytime sleepiness</td>
<td>7–12 years</td>
<td>18 items; child report; 1 week retrospective survey</td>
<td>Total: $\alpha = .71$</td>
<td>Correlated with CSHQ</td>
<td>Approaching well-established</td>
</tr>
</tbody>
</table>
consists largely of concurrent validity with other sleep habits measures. The reliability of measures ranged from acceptable (Family Inventory of Sleep Habits) to good (ASHS, BRQ, CSHS). The Adolescent Sleep Hygiene Scale received “approaching well-established” EBA classification, and the Family Inventory of Sleep Habits, the BRQ, and the CSHS were rated as “promising”.

The CSHS is a parent-report measure of activities surrounding sleep in children aged 2–8 years. The scale was rated as “promising” because the measure description and psychometric information were published exclusively in abstract form rather than in a full-length article. The measure was subsequently used by other investigators in studies examining sleep hygiene in children (Henderson & Jordan, 2010; van der Heijden, Smits, & Gunning, 2006). In terms of validity, the CSHS shows concurrent validity with another measure of sleep habits, the BRQ.

The ASHS is an adolescent-report measure designed to assess behaviors that may inhibit or facilitate sleep in adolescents (e.g., nighttime caffeine intake, level of activity before bedtime). The ASHS was modified from the CSHS. It was rated as “approaching well-established” because while it has been presented in at least two peer-reviewed journal articles, validation information is only moderate. Regarding validity, the ASHS shows concurrent validity with associations with the Adolescent Sleep Wake Scale (ASWS) (LeBourgeois, et al., 2005). The ASHS is the only measure in this category appropriate for use with youth older than 12 years, and is the only measure developed for child or adolescent self-report.

The BRQ is a parent-report questionnaire designed to assess sleep habits/hygiene in children aged 2–8 years. Of all the measures in this category, the BRQ has the most published support for measure validity. A factor analysis of the BRQ was conducted and subscale items correlated with the Children’s Sleep Hygiene Scale (CSHS) and Children’s Sleep Wake Scale (CSWS). In addition, scores on the BRQ differentiated between good and poor sleepers (Henderson & Jordan, 2010). Despite this validation data, the BRQ received a “promising” classification primarily because it has not yet been used by other investigative teams.

The Family Inventory of Sleep Habits (Malow et al., 2009) is a unique sleep assessment tool because it was developed for children with autism. Of all the measures reviewed, this was the only measure developed for a specific clinical or psychiatric population. Some questionnaire items do not directly assess sleep hygiene, but rather overall sleep behaviors relevant to children with autism (e.g., sleeping in pajamas made of certain fabrics, sleeping with a comfort object). Concurrent validity of the Family Inventory of Sleep Habits is demonstrated through correlations with the Children’s Sleep Habits Questionnaire (CSHQ) (Owens, Spirito, & McGuinn, 2000). Currently rated as “promising,” this measure needs additional validation and use by other researchers.

**Measures of sleep initiation, sleep maintenance, and sleep quality**

Five measures evaluated sleep initiation, maintenance, and quality: the ASWS (LeBourgeois et al., 2005), the CSWS (LeBourgeois & Harsh, 2001), the Infant Sleep Questionnaire (ISQ) (Morrell, 1999a), the Sleep and Settle Questionnaire (SSQ) (Matthey, 2001), and the Tayside Children’s Sleep Questionnaire (TCSQ) (McGreavey, Donnan, Pagliari, & Sullivan, 2005). The assessment tools in this category assess problems initiating and maintaining sleep; however, they vary in their target population and scope. All of the measures in this category have moderate to high internal consistency and/or test–retest reliability, and show some support for validity. The ISQ was the only measure to receive a “well-established” EBA classification. The Tayside (McGreavey et al., 2005) and the ASWS received “approaching well-established” classifications and the SSQ (Matthey, 2001) and CSWS were rated as “promising.”

The ISQ (Morrell, 1999a) and the SSQ (Matthey, 2001) are parent-completed measures designed to screen for difficulties settling and nightwakings in infants. The ISQ was rated as “well-established” because it has good reliability data, sensitivity and specificity for identifying sleep problems, and has been used by different investigative teams (e.g., Morrell, 1999b; Schuetze, Lawton, & Eiden, 2006). The SSQ was rated as “promising” because it has not been used by researchers other than the authors who developed the measure. Psychometric properties of the SSQ are good, with the measure showing sensitivity to change/treatment effects, and the ability to differentiate between mothers who did/did not report sleep problems in their infants (Matthey, 2001).

The Tayside (McGreavey et al., 2005) assesses prevalence and severity of problems initiating and maintaining sleep for children aged 1–5 years. The Tayside was rated as “approaching well-established” because although the original publication presents strong support for reliability and validity (McGreavey et al., 2005), it has thus far been used by only one other investigative team (Johnson & McMahon, 2008). Greater use of this measure is needed to expand psychometric data and provide more support for its use.

The CSWS (LeBourgeois & Harsh, 2001) and the ASWS (LeBourgeois et al., 2005) both examine five
behavioral sleep dimensions (falling asleep, awakening, maintaining and reinitiating sleep, returning to wakefulness) with the CSWS used for children aged 2–8 years, and the ASWS in adolescents aged 12–18 years. The CSWS was also rated as promising because although it has been used in a peer-reviewed study examining sleep quality in children with ADHD (LeBourgeois, Avis, Mixon, Olmi, & Harsh, 2004), the measure and corresponding validation data were published exclusively in abstract form (LeBourgeois, Hancock, & Harsh, 2001; LeBourgeois & Harsh, 2001). Similarly, while the ASWS has some early psychometric support (e.g., good reliability, concurrent validity with another adolescent sleep measure) and has been used by other investigative teams (Palermo, Fonareva, & Janosy, 2008), it was rated as “approaching well-established” because additional validation data are needed.

**Measures of daytime sleepiness**

Four measures assessed daytime sleepiness behaviors such as falling asleep and alertness during the day. The Pediatric Daytime Sleepiness Scale (PDSS) (Drake et al., 2003) was the sole measure of daytime sleepiness rated as “well-established.” The modified version of the Epworth Sleepiness Scale (ESS) (Melendres, Lutz, Rubin, & Marcus, 2004) and the Teacher’s Daytime Sleepiness Questionnaire (TDSQ) (Owens et al., 1999) were rated as “approaching well-established,” and the Cleveland Adolescent Sleep Questionnaire (CASQ) (Spilsbury, Drotar, Rosen, & Redline, 2007) was classified as “promising.”

Originally an adult measure, the ESS was modified by investigators to make it more appropriate for use with adolescents (e.g., sleepiness while “driving in traffic” was replaced with “doing homework or taking a test”) (Melendres et al., 2004; Moore et al., 2009). Studies using the modified ESS have shown preliminary support for validation; adolescents with sleep disordered breathing scored higher on the ESS than controls (Melendres et al., 2004) and scores on the ESS correlated with self-reports of anxiety and general health status (Moore et al., 2009). The ESS has also shown weak correlations with polysomnography (Melendres et al., 2004). While the modified ESS has good initial support, the measure was rated as “approaching well-established” because individual items on the modified ESS differed across studies. Greater validation of the modified version is needed to provide further support for use with adolescents.

The TDSQ is a teacher-report questionnaire that assesses classroom behaviors of children (aged 4–10 years) that are likely associated with nighttime sleep disturbance. Validation studies have demonstrated significant but low magnitude correlations with daytime sleepiness on the CSHQ (Owens, Spirito, McGuinn et al., 2000) and a factor analysis was conducted (Owens et al., 1999). The TDSQ were classified as “approaching well-established.” The measure has preliminary psychometric support; however, more data are needed such as test–retest reliability and association with other sleep measures.

The CASQ is a measure of sleepiness developed for adolescents (11–17 years). The CASQ received a rating of “promising” because although it has high internal consistency and convergent validity with other sleep measures (PDSS and the Sleep Habits Survey), it has thus far only been used by one investigatory team. In terms of additional psychometric support, the CASQ has been shown to detect differences in daytime sleepiness in a clinical sample of youth with obstructive sleep apnea and healthy controls (Spilsbury et al., 2007).

The final measure in this category, the PDSS, is brief 8-item scale that assesses daytime sleepiness and includes specific questions about impact on academic performance (Drake et al., 2003). The PDSS received a “well-established” rating because it has excellent psychometric properties and has been used by several investigators. This measure has been used with clinical samples (e.g., obesity and obstructive sleep apnea) and reliability data for the Spanish version is available (Beebe et al., 2007; Perez-Chada et al., 2007).

**Measures of sleep-related beliefs and cognitions**

Two measures evaluated sleep-related beliefs and cognitions, the Dysfunctional Beliefs about Sleep Questionnaire (DBAS) (Gregory, Cox, Crawford, Holland, & Harvey, 2009; Morin, Stone, Trinkle, Mercer, & Rensberg, 1993), and the Presleep Arousal Survey for Children (PSAS-C) (Gregory, Willis, Wiggs, & Harvey, 2008). The DBAS and PSAS-C were originally designed and validated for use with adults. Both measures were revised to become child-report assessment tools that examine children’s dysfunctional thoughts about their sleep (e.g., how much sleep children feel they need, causal attributions of insomnia). The DBAS and the PSAS-C are relatively short questionnaires (24 and 16 items, respectively) and both include subscale and total scores.

In addition to cognitions about sleep, the PSAS-C measures somatic arousal (e.g., rapid pulse, sweating palms) prior to sleep. In terms of psychometric properties, this measure has adequate internal consistency and the measure has been associated with sleep disturbances on the CSHQ and the Sleep Self Report (SSR) (Gregory et al., 2008). Recently, the PSAS-C has been used with older children and adolescents and has demonstrated reliability and
concurrent validity for adolescents up to 18 years (Alfano, Pina, Zerr, & Villalta, 2010). The PSAS-C was rated as “approaching well-established” because while it has been presented in at least two peer-reviewed journal articles, validation information is only moderate and needs to be expanded.

The DBAS was rated as “promising” due to its limited psychometric data for use with pediatric populations and the wide range of subscale internal consistency. Only one article has been published presenting psychometric data on use of the measure with children. These investigators found that the DBAS was associated with children’s sleep disturbance (as assessed by the CSHQ) (Gregory et al., 2009).

Multidimensional pediatric sleep measures

Six multidimensional sleep measures were reviewed with four of these multidimensional measures receiving “well-established” EBA classifications. The multidimensional sleep measures vary significantly in terms of length and population targeted, and the measures screen for a broad range of sleep problems such as sleep habits/hygiene, daytime sleepiness, parasomnias, and nightwakings. The four measures rated as “well-established” were the Brief Infant Sleep Questionnaire (BISQ) (Sadeh, 2004), the Children’s Sleep Habits Questionnaire (CSHQ), (Owens, Spirito, & McGuinn, 2000), the Pediatric Sleep Questionnaire (PSQ) (Chervin, Dillon, Bassett, Ganoczy, & Pituch, 1997; Chervin, Hedger, Dillon, & Pituch, 2000), and the Sleep Disturbance Scale for Children (SDSC) (Bruni et al., 1996). The two measures that received “approaching well-established” classifications were the Sleep Self-Report (SSR) (Owens, Maxim, Nobile, McGuinn, & Msall, 2000) and the Sleep Habits Survey (SHS) (Wolfson & Carskadon, 1998; Wolfson et al., 2003).

The BISQ is the only multidimensional measure appropriate for use with infants. It is a short 13-item parent-report screening tool that assesses sleep problems in infants aged 0–29 months. In contrast, the CSHQ and SSR are designed for school-age children (4–10 years and 7–12 years, respectively), and the PSQ, SDSC, and SHS are designed for use from childhood through adolescence. The PSQ has the widest age range and can be used with youth aged 2–18 years; moreover, the subjective sleepiness subscale of the PSQ is the only scale reviewed that has been validated with an objective sleepiness measure (Chervin et al., 2000). In addition to behaviorally based sleep questions, the PSQ, SDSC, and CSHQ include questions that target sleep-related breathing disorder, snoring, parasomnias and other clinical sleep problems (e.g., narcolepsy, nightmares, enuresis). The CSHQ, PSQ, SDSC, and SHS also provide subscale scores for examining specific sleep disturbances.

Because of their strong psychometric support and use by multiple investigators, the BISQ, CSHQ, PSQ, and the SDSC were rated as “well-established”. The SSR and SHS were rated as “approaching well-established” because they had not been used as widely and have more limited validity data (e.g. factor analysis, differentiate between clinical and non-clinical populations). With respect to overall psychometric properties, all six of the measures present test–retest data for the total scale or individual items that range from satisfactory to excellent. Internal consistency statistics were available for all measures except the BISQ (total or subscale scores) with values ranging from moderate to good. The CSHQ and the SDSC have separate internal consistency scores for clinical and nonclinical populations.

In terms of validity, the BISQ has been correlated with actigraphy and daily sleep logs, and shows sensitivity to developmental sleep changes (Sadeh, 2004). Similarly, scores on the SHS were correlated with both prospective diary and actigraphy variables (Wolfson et al., 2003). The SDSC and the CSHQ both differentiate between clinical and control groups, with the SDSC showing good diagnostic accuracy (AUC = .91) (Bruni et al., 1996; Owens, Spirito, & McGuinn, 2000). Both the PSQ and SDSC have been factor analyzed.

Strengths of measures reviewed

Of the 21 measures included in the current review, six met criteria for a “well-established” classification, indicating that some sleep assessment tools have a strong evidence base for use with children and adolescents. Eight measures met criteria for “approaching well-established” and seven were categorized as “promising.” Across categories, many measures present good reliability and validity data. Nine of the measures (or their subscales) were able to differentiate between healthy youth and those with sleep disorders (e.g., BRQ, CASQ, CSHQ, ESS, PSAS-C, PSQ, SDSC, SSQ, TDSQ). About half of the questionnaires have been used with medical or psychiatric populations and demonstrated reliability/validity in these clinical samples. For example, the PSQ, PDSS, and SDSC were used with children with epilepsy, and the ASWS and the SSR were used with pain populations. The CSHQ has been used with the widest variety of populations (e.g., autism spectrum disorders, mental retardation, obesity, ADHD), followed by the SDSC. The Family Inventory of Sleep Habits (developed for with children with autism) and was the only condition-specific measure reviewed.

Many of the measures reviewed (ASWS, ASHS, BISQ, CSHQ, PDSS, PSQ, SDSC, SHS, SSR) and their
corresponding subscales have been translated into different languages (e.g., Spanish, Chinese, Italian). Notably, the BISQ, CSHQ, PDSS, PSQ, and SDSC have multiple translations. Recently researchers have been comparing sleep problems in youth from different cultures, showing both similarities and differences between American youth and children in Europe and Asia (LeBourgeois et al., 2005; Liu, Liu, Owens, & Kaplan, 2005).

An additional strength is that in the validation process, some of the measures were validated against other forms of sleep assessment. For example, in validation studies of the BISQ and Sleep Habits Survey, associations were demonstrated with actigraphy (Sadeh, 2004; Wolfson et al., 2003). Similarly both the PSQ and SDSC’s sleep-related breathing disorder subscales have been validated with polysomnographic-confirmed sleep-related breathing disorders (Chervin et al., 2000; Ferreira et al., 2009). These validation studies are useful for understanding the relationship between questionnaire reports and other methods of sleep assessment. In addition to studies aimed at measure validation, many other measures (e.g., ASWS, ASHS, CSHQ, ESS, PDSS, PSAS-C, PSQ, Sleep Self-Report) have been used in studies comparing these questionnaires with other sleep assessment tools (e.g., Beebe et al., 2007; Carno et al., 2008; Chervin et al., 2007; El-Sheikh, Hinnant, Kelly, & Erath, 2010; Kim et al., 2010; Moore et al., 2009; Ward et al., 2008). Such research provides further support for measure use.

**Limitations of measures reviewed and future directions**

The purpose of this review was to provide an evidence-based evaluation of parent and child-report sleep measures used by psychologists with children and adolescents. Six of the sleep measures reviewed received a “well-established” rating and many other measures show promise as potentially useful tools for research and clinical assessment of sleep in youth. However, there are weaknesses evident in the measures as a whole. Further development of sleep assessment tools in the areas listed below are considered current research priorities.

1. Many of the measures reviewed provided limited psychometric data and additional information on reliability and validity of existing measures is needed. Only seven of the measures reported test–retest reliability values. Moreover, although many measures presented information on convergent validity, data on stability over time and predictive, concurrent, and construct validity are lacking.

2. Use of measures in longitudinal studies to determine how they function in response to maturation and treatment are also needed. After their initial validation, the minority of measures (e.g., CSHQ, PSQ) were subsequently used in longitudinal research (e.g., Chervin, Ruzicka, Archbold, & Dillon, 2005; Gregory, Rijsdijk, Dahl, McGuffin, & Eley, 2006). Identifying measures that are sensitive to longitudinal changes in sleep will be particularly important in the future as focus shifts to assessment of treatment outcomes.

3. Factor analysis is still needed for most measures to better understand the latent structure and dimensions of the constructs being assessed. At present, factor analysis has been conducted on only seven of the measures reviewed. Additional factor analytic data can be used to better characterize sleep domains and aid in interpretation of sleep measures.

4. The minority of sleep measures reviewed underwent validation with other types of sleep assessment tools (e.g., actigraphy, polysomnography). Further studies aimed at validation with such tools will provide greater psychometric support for use of sleep questionnaires, and inform researchers on how assessment tools differentially capture sleep disturbances.

5. The use of technology (e.g., web and computer administration) is also a key future direction. Use of technology will also allow clinicians and researchers to more easily track changes in sleep patterns/behaviors over time without requiring youth to come in to the clinic or office.

6. Obtaining the measure for use in a research or a clinical setting is also a limitation. Although some of the measures and their scoring instructions are accessible on the authors’ websites or online (e.g., CSHQ, PSQ, Sleep Habits Survey), other measures can only be obtained by contacting the author directly (e.g., ASHS, ASWS). The difficulty in obtaining some of these questionnaires may be a barrier in their use.

7. Development of new measures and/or modification of current questionnaires to permit cross-informant comparisons are also needed. Only one measure reviewed (modified ESS) has both parent and child versions making cross-informant comparisons about sleep perceptions possible. While for certain measures (e.g., BISQ, SSQ) child report would not be possible due to age, sleep assessment with
multiple adult informants (parents, teachers) could be conducted. Differences in reporter accuracy when assessing children’s externalizing and internalizing behaviors (e.g., Loeber, Green, Lahey, & Stouthamer-Loeber, 1991) support the importance of examining such differences on reports of children’s sleep.

8. The majority of the measures reviewed failed to take into account the need for different types of items and response formats for youth of different ages and developmental levels. The only behavioral sleep measures with different versions (and published reliability and validity data) for children and for adolescents were the CSHS and ASHS and the CSWS and the ASWS. Using the same items to assess sleep for these disparate ages fails to consider the potential developmental differences in sleep patterns and behaviors. In addition, most measures assess only problem frequency but other perceptions are important to consider. For example, the CSHQ asks not only about the frequency of sleep behaviors, but also whether or not the parent considers these behaviors to be a problem. Modification of existing measures or development of new measures is needed to address this issue.

9. The time frame for many measures reviewed varies widely (e.g., present, past week, past 3 months) and can be vague (e.g., CSHQ—report on past week, or past typical week). These differences in scales can make comparisons across measures challenging. In addition by not differentiating between weeknights and weekends, it can be difficult for youth to accurately report on a “typical” night’s sleep.

10. The minority of the measures (PDSS, Sleep Habits Survey, TDSQ) were developed in a manner sensitive to the varying social contexts affected by children’s sleep (e.g., home and school). Recognizing strong associations among sleep and psychological and socioemotional outcomes, it is important to expand this assessment of sleep into other domains. Given the research showing that daytime sleepiness impacts school performance and neurobehavioral functioning (Sadeh et al., 2002, 2003), validation of existing or creation of new measures that assess impact of sleep on school are needed.

11. This review also revealed that for certain sleep domains, the number of available measures that are reliable and valid for youth of different ages are limited. For example, while many of the multidimensional measures include assessment of sleep habits/hygiene, there are no “well-established” sleep habits-specific measures and the only tool in this category appropriate for use with adolescents was the ASHS. While use of multidimensional measures to assess these constructs may be appropriate in clinical settings, response-burden associated with the use of longer multidimensional questionnaires may be a barrier, particularly epidemiological research studies.

12. Importantly, insomnia-specific questionnaires are lacking in pediatric sleep assessment. While there are validated adult insomnia measures (e.g., The Insomnia Symptoms Questionnaire; Okun et al., 2009), we were unable to locate any measures designed specifically to assess insomnia symptoms in children and adolescents. Given the impact of insomnia on this population, the creation of new questionnaires in this area should be considered a research priority.

13. Finally, validation of existing sleep measures with more diverse populations is needed. While some studies do report validation with racially heterogeneous samples (e.g., BRQ, DBAS, PSAS, CASQ) or include multiple translations (ASWS, ASHS, BISQ, CHSQ, PDSS, PSQ, SDSC, SHS, SSR), additional work is needed. The failure to include diverse samples in measure development limits generalizability of current measures. Future studies with more diverse populations will help establish if measures function similarly across cultures or ethnic groups.

Acknowledgment
The authors would like to thank Erika Doty and Rachel Higginbotham, M.D. MPH for their assistance with this project.

Funding
This work was supported by the National Institute of Health (RO1HD053431 and K24HD060068 to T.M.P.).

Conflicts of interest: None declared.

References
monitoring in children and adolescents: How many nights are necessary for reliable measures? Sleep, 22(1), 95–103.


