Evidence-based Assessment of Pediatric Pain

Lindsey L. Cohen,[1] PhD, Kathleen Lemanek,[2] PhD, Ronald L. Blount,[3] PhD, Lynnda M. Dahlquist,[4] PhD, Crystal S. Lim,[1] MA, Tonya M. Palermo,[5] PhD, Kristine D. McKenna,[4] MS and Karen E. Weiss,[4] MS

[1]Georgia State University, [2]College of Medicine, Ohio State University, [3]University of Georgia, [4]University of Maryland Baltimore County, and [5]Oregon Health and Science University

Objective To conduct an evidence-based review of pediatric pain measures. **Methods** Seventeen measures were examined, spanning pain intensity self-report, questionnaires and diaries, and behavioral observations. Measures were classified as ''Well-established,'' ''Approaching well-established,'' or ''Promising'' according to established criteria. Information was highlighted to help professionals evaluate the instruments for particular purposes (e.g., research, clinical work). **Results** Eleven measures met criteria for ''Well-established,'' six ''Approaching well-established,'' and zero were classified as ''Promising.'' **Conclusions** There are a number of strong measures for assessing children's pain, which allows professionals options to meet their particular needs. Future directions in pain assessment are identified, such as highlighting culture and the impact of pain on functioning. This review examines the research and characteristics of some of the commonly used pain tools in hopes that the reader will be able to use this evidence-based approach and the information in future selection of assessment devices for pediatric pain.

Key words assessment; interviews; observational; pain; pediatrics; self-report.

Pain is the most common reason people present for health care, pain costs to society are exorbitant, and pain can have a widespread impact on all aspects of life (Stewart, Ricci, Chee, Morganstein, & Lipton, 2003). The importance of attending to pain is highlighted by the Joint Commission on Accreditation of Healthcare Organizations (JCAHO; Phillips, 2000) directives that it be considered the ''fifth vital sign'' to monitor in medical care. However, it is not readily evident how pain should be monitored.

Pain is both a sensory and emotional personal experience, making assessment complex (Melzack & Wall, 1965). Further, pain occurs across a spectrum of conditions including acute injuries and medical events, recurrent or chronic pain, and pain related to chronic disease. Acute pain is typically brief, ending around the time of the healing of an injury, or the termination of the stretching, contraction, or impingement of some part of the body (Cohen, MacLaren, & Lim, 2007). Chronic pain, on the other hand, may or may not be symptomatic of underlying, ongoing tissue damage or chronic disease. It can persist long after an initial injury has healed or other event has occurred (typically longer than 3 months) (Cohen, MacLaren, & Lim, 2007). Clearly, evaluating pain is complicated both by the personal nature of the experience and the variety of forms in which it can exist.

The assessment of children's pain is especially problematic as younger children or those with developmental delays often do not have the language or cognitive sophistication to describe their pain. Unfortunately, pain is a frequent and vivid part of childhood, whether as part of routine care (e.g., immunization injections) or a symptom of a chronic illness (e.g., chronic sickle cell disease pain). Outside of the medical arena, children experience frequent bumps, bruises, and injuries as they acquire coordination and adapt to their quickly developing body.

Accurate assessment of children's pain is needed to diagnose conditions and to guide pain management interventions; especially given the accumulating research

All correspondence concerning this article should be addressed to: Lindsey L. Cohen, PhD, Department of Psychology, Georgia State University, Atlanta, GA 30302-5010. E-mail: llcohen@gsu.edu

Journal of Pediatric Psychology 33(9) pp. 939–955, 2008
doi:10.1093/jpepsy/jsm103
Advance Access publication November 17, 2007

suggesting that untreated pain may have long-term negative and permanent repercussions on pain sensitivity, immune functioning, neurophysiology, attitudes, and health care behavior (for a review, see Young, 2005). Instruments measuring pain intensity, location, and affect are typically used to assess acute pain of relatively brief duration. Measurement of recurrent and chronic pain requires tools that also measure the frequency, duration, time course, and activity interference due to pain.

Over the past 15 years, significant research attention has been devoted to developing instruments to quantify children's pain. Whereas there are descriptions and summaries of measures (Finley & McGrath, 1998; O'Rourke, 2004), there are only a few critical evaluations and comparisons of pain measures. Recently, the Pediatric Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (Ped-IMMPACT) commissioned reviews to identify measures to use in pediatric pain clinical research trials. The Ped-IMMPACT papers have included one on self-report (Stinson, Kavanagh, Yamada, Gill, & Stevens, 2006) and one on observational measures (von Baeyer & Spagrud, 2007). These two reviews and the current report have similarities, for example, both used the same Society of Pediatric Psychology Assessment Task Force criteria (Cohen, La Greca, Blount, et al., 2007). In addition, where the groups overlapped, there are generally consistent recommendations. For example, all measures that were included in both the Ped-IMMPACT and current reports were rated as ''well-established.''

Although there are consistencies, the current review and the Ped-IMMPACT reviews differ in a number of ways. One distinction is that the current review spans a broad focus on pain intensity, distress behaviors, and caregiver behaviors using self-report, questionnaire, diary, and observational instruments, as well as infants through adolescents. The Ped-IMMPACT reviews focus on children and adolescents 3–18 years of age and focus on self-report (Stinson) and observational (von Baeyer & Spagrud) measures. The most distinguishing aspect of the current review is its purpose. Specifically, we sought to apply an evidence-based framework for evaluating a number of popular and diverse pain tools that are used by pediatric psychologists. We present the strengths and limitations of measures used in the field of pediatric psychology to assess pain. Thus, interested parties can make informed decisions in critiquing and selecting measures for their particular purpose, population, and situation (e.g., clinical work, randomized clinical trials). In other words, whereas the Ped-IMMPACT reviews take a more nomothetic approach to establish broad recommendations or guidelines for clinical trials, the current review adopts an idiographic perspective to help professionals identify measures to answer their unique questions. Given that both groups relied on a combination of objective and systematic as well as expert opinion or subjective methods in the selection and evaluation of measures, there were measures distinct to Ped-IMMPACT or the current review. For example, the current review did not include the FLACC (Merkel, Voepel-Lewis, Shayevitz, & Malviya, 1997) or the Post-operative Pain at Home Scale (PPPM; Chambers, Reid, McGrath, & Finley, 1996), which were included by von Baeyer & Spagrud. Likewise, a number of measures included in the current review were omitted in the Ped-IMMPACT papers. The various differences between the Ped-IMMPACT and this review may be attributed to the differences in the purpose, process, methodology, and decisions of the reviewing committees.

Framework

The American Psychological Association (APA), Division 54, Society of Pediatric Psychology (SPP) Assessment Task Force commissioned this review to provide evidence-based assessment of pediatric pain. We adopted the APA definition of evidence-based practice, which states that ''Evidence-based practice in psychology is the integration of the best available research with clinical expertise in the context of patient characteristics, culture, and preferences'' (APA Presidential Task Force on Evidence-based Practice, 2006). For the purposes of our review, along with ''patient,'' we have included ''research participant,'' ''research sample,'' ''setting,'' and other relevant situational variables. From this approach, we acknowledge that selection of a pain measure will depend on a number of factors including the research base (e.g., psychometrics, prior findings), patient/situation characteristics (e.g., age, culture, time available for assessment), and clinical expertise (e.g., weighing particular research findings and situational aspects; experience needed to perform, score, or interpret the measure). In our review of measures, we highlight some of the research base and relevant patient/situation characteristics. Where pertinent, we identify aspects of the measure related to needed expertise.

Measure Selection

Selection of measures was conducted in multiple stages. First, in 2002, a list of assessment measures was generated by the SPP Assessment Task Force in broad areas of interest, one of which was pain. Second, the survey was distributed via the Internet to the SPP listserv. Eighty-seven respondents completed the survey.

These respondents identified an additional three pain measures not initially included in the survey. Third, we discussed the responses, surveyed the literature, and consulted other pain measure review sources. This process occurred during 2005–2006. Based on this search and evaluation, we selected a pool of measures seen as those that are commonly used by pediatric psychologists across self-report, interviews and questionnaire, observational, and diary formats. We selected instruments that represented a range of type of measure (e.g., self-report, observational) and are popular in the pediatric psychology literature. In 2007, another literature search was conducted specifically to identify additional studies using the measures included in our review. This review was not intended to be exhaustive or comprehensive. Given that the best measures will depend on the questions being asked, and that the pain assessment research literature is evergrowing; we did not attempt to detail a list of the ''best'' measures, but rather, we adopted an evidence-based approach in evaluating popular pediatric pain measures. The final list of 17 measures consisted of five pain intensity self-reports, four questionnaire and diary, and eight observational instruments (Table I). When making the decision of which measures to highlight in our review, we aimed for presenting a range of measures and ones that are commonly used by pediatric psychologists. We acknowledge that the review is not comprehensive, and that the available research base for pain measure is ever-changing.

Assessment Criteria for Research Base

As detailed in the lead article in this series, assessment criteria were developed to apply to the measures reviewed (Cohen, La Greca, Blount, et al., 2007), in order to establish some of the research base supporting the measure. Accordingly, the instruments were classified as ''Well-established assessment'' (e.g., at least two research teams have published, sufficient information available to evaluate the measure, good psychometric properties), ''Approaching well-established'' (e.g., presented in at least two articles, sufficient detail available, moderate or vague psychometrics presented), or ''Promising assessment'' (e.g., at least one peer-reviewed article, sufficient detail available, moderate or vague psychometrics). Although interpretation and application of these criteria were left to the workgroup authors, the SPP Assessment Task Force circulated documents detailing considerations for interpreting the validity and reliability of a variety of measures. Inter-rater agreement was assessed to evaluate the reliability of our ratings of the pain measures. Specifically, a detailed summary of each measure was developed and was sent along with the task force criteria to a pain researcher who was not an author or collaborator on the current review. The authors' ratings and those of the outside rater showed high agreement with only one instance of disagreement, which resulted in a κ of .87 and a weighted κ of .89.

We should note that the criteria used by this committee highlight the quality of the research base for a given measure, but do not provide an evaluation of the overall quality of a measure. In order to truly critique a measure and deem it the best one in a particular situation, it is necessary to know the questions being asked, the particular personal and contextual circumstances in which it will be used, and the expertise of the individual administering the instrument.

Review of Measures

Pain Intensity Self-Report

In adults, it has been recommended that pain intensity be assessed by asking patients to rate their pain on a numerical rating scale, with 0 indicating no pain and 10 indicating the worst pain possible (Dworkin et al., 2005). Because of children's more limited understanding of number concepts, a variety of other rating scales have been developed in which children provide a graphic representation of pain intensity by marking a point on a line (VAS), pointing to or coloring a certain level on a pain ''thermometer,'' selecting from a series of faces depicting different levels of pain, or counting different numbers of simple, tangible objects, such as poker chips that allow the child to indicate more or less pain.

VAS typically consist of a 100 mm horizontal line with anchors indicating ''no pain'' at the left endpoint and ''worst pain possible'' (or a comparable term) at the right endpoint. The child makes a vertical mark on the line to indicate how much pain he/she feels. Pain intensity scores are calculated by measuring the distance from the left end point of the scale to the child's mark. The exact wording of anchors varies from study to study. Examples include: ''no pain'' versus ''very severe pain,'' ''worst pain ever,'' ''extreme pain,'' ''pain as bad as it could be,'' and ''very much pain'' (Hicks, von Baeyer, Spafford, van Korlaar, & Goodenough, 2001; Polkki, Pietila, & Vehvilainen-Julkunen, 2003).

Extensive evidence supports VAS pain ratings as valid indicators of children's pain experience. Children's VAS scores have been shown to correlate significantly with parent ratings of children's pain (Luffy & Grove, 2003; Varni, Thompson, & Hanson, 1987) and with medical

Table 1. Reviewed Pediatric Pain Measures

Measure	Brief Description	Ages (years)	Psychometrics	EBA Rating
Pain intensity self-report				
Visual analog scale (VAS)	Self-report visual analog scales for pain intensity. Horizontal line with descriptive pain anchors at endpoints; draws line that intersects to indicate intensity.	3–adult	Inter-rater correlations = .28–.72[a] Concurrent validity = .61–.90 Test-retest reliability = .41–.58	Well-established
The Oucher (Beyer, 1984)	Self-report photograph scale for pain intensity.	3–12	Concurrent validity = .62–.95 Test-retest reliability = 78% of children reported scores within ± one level after 15 min.	Well-established
Wong-Baker Faces Pain Rating Scale (Wong & Baker, 1988)	Self-report faces scale for acute pain. six line drawn faces range from no hurt to hurts worst.	3–18	Concurrent validity: Other pain measures = .67–73 Inter-rater correlations = .26–.37	Approaching well-established
Faces Pain Scale-Revised (FPS-R; Hicks, von Baeyer, Spafford, van Korlaar & Goodenough, 1993)	Self-report faces scale for acute pain. six cartoon faces range from neutral to high pain expression.	4–16	Concurrent validity = .84–.99 Inter-rater correlations = .84–.99	Well-established
Poker chip tool (Hester, 1979)	Self-report poker chips are used to represent pain intensity. Child chooses which chips represent the pain they experience	4–7	Inter-rater correlations = .23–.70 Concurrent validity = .65–.94 Test-retest reliability (8 hr) = .83	Well-established
Questionnaire and Diaries				
Headache Diary (Richardson, McGrath, Cunningham, & Humphreys, 1983)	Likert scale used to assess intensity of headache pain four times a day. Scale can be behavioral or subjective	8–17	Concordance (weighted κ) = .18–1.0	Approaching well-established
Pain Diary (Hunfeld et al., 2001; Hunfeld et al., 2002)	Visual analogue scale assessing intensity of current pain three times a day	12–18 (parent proxy 5–11)	Test-retest reliability = .88–.98 Concurrent validity = .46–.84 Convergent validity = .03–.56	Approaching well-established
Abu-Saad Pediatric Pain Assessment Tool (PPAT; Abu-Saad, Kroonen, & Halfens, 1990)	Self-report multidimensional questionnaire of pain using 32 sensory, affective, and evaluative word descriptors, and a 10 cm scale that measures present and worst pain.	5–17	Internal consistency = .83 Cross-informant correlations = .28–.92 for child, parent, nurse, physician present and worse pain Concurrent validity = .94–.97 (10 cm); .45–.51 (NWD) other pain measures; Ns 10 cm and Headache Diary Convergent validity = .50–4.50 Discriminant validity = .14–.26	Approaching well-established
Varni-Thompson Pediatric Pain Questionnaire (PPQ; Varni & Thompson, 1985)	Questionnaire that assesses chronic pain intensity, location, sensory, evaluative, and affective qualities of pain via self-report and parent / physician proxy-report. Used with a variety of populations (e.g., JRA, SCD, fibromyalgia).	5–18	Test-Retest reliability = .29–.41 Inter-rater correlations = .40–.85 VAS predictive of disability estimates ($p < .05$) Convergent Validity = .27–.68 with disease status; .06–.45 with psychological functioning	Well-established

(continued)

Table 1. Continued

Measure	Brief Description	Ages (years)	Psychometrics	EBA Rating
Behavioral Observation Procedure Behavioral Rating Scale (PBRS; Katz, Kellerman, & Siegel, 1980)	Observational measure of behavioral distress (pain, anxiety, and fear) during painful medical procedures. Thirteen operationally defined behaviors.	0–18	Inter-rater reliability = .81–.94 Convergent validity = .33–.68	Approaching well-established
Observational Scale of Behavioral Distress (OSBD; Jay, Ozolins, Elliott, & Caldwell, 1983; OSBD-R, Elliott, Jay, & Woody, 1987)	Observational measure of pain during acute medical procedures. Originally consisted of 11 behaviors that indicate distress, but has been revised to include eight.	2–20	Internal consistency = .68–.72 Concurrent validity = .20–.76 Inter-rater reliability = 80–91%	Well-established
Child–Adult Medical Procedure Interaction Scale (CAMPIS; Blount et al., 1989; CAMPIS-R; Blount et al., 1997)	Observational measure used to assess child pain during acute medical procedures. Child coping and parent / medical staff behaviors are also assessed.	2–13	Inter-rater reliability = .65–.92 for the different scales; .90 for distress (κ)	Well-established
Procedure Behavior Checklist (PBCL; LeBaron, & Zeltzer, 1984)	Observational measure of pain and anxiety during invasive medical procedures. Eight operationally defined behaviors rated on occurrence and intensity (scale 1–5)	3–18	Inter-rater reliability = 72–94% Convergent validity = .42–.74	Well-established
Children's Hospital of Eastern Ontario Pain Scale (CHEOPS; McGrath, Johnson, Goodman, Dunn, & Chapman, 1985)	Observational measure of postoperative pain in children.	1–12	Inter-rater reliability = 95% Concurrent validity = .35–.85	Well-established
COMFORT Scale (Ambuel, Hamlett, Marx, & Blumer, 1992)	Observer rated measure for use in intensive care environments. Includes eight behaviorally anchored rating scales.	0–18	Inter-rater reliability = .51–.93 (κ) Concurrent validity = .26–.90 Internal consistency = .90–.92	Well-established
Child Facial Coding System (CFCS; Chambers, Cassidy, McGrath, Gilbert, Craig, 1996).	Observational measure of facial expressions during painful medical procedures. Thirteen facial actions coded for frequency or intensity	2–5	Inter-rater reliability = .75–.83 Concurrent validity = .28–.73	Approaching well-established
Premature Infant Pain Profile (PIPP; Stevens, Johnston, Petryshen, & Taddio, 1996)	Observational measure of acute pain in premature infants. Seven indicators of pain	Preterm infants	Internal consistency = .59–.76 Inter-rater reliability = .95–.97	Well-established

Note: *The psychometrics are from a number of sources found via psycINFO and MEDLINE searches conducted between 2003–2007. For an overview of the measure including the relevant references, please contact the corresponding author.

personnel ratings (Gragg et al., 1996). VAS scores correlate positively with scores obtained on other pain scales, such as the Oucher (Aradine, Beyer, & Thompson, 1988; Beyer & Aradine, 1987), the Eland Color Scale (Guariso et al., 1990), various faces scales (Hunfeld, van der Wouden, den Deurwaarder, van Suijlekom-Smit, & Hazebroek-Kampschreur, 1997), and the COMFORT Scale (van Dijk et al., 2000). VAS scores correlate positively with disease indicators commonly associated with pain, such as joint swelling and disease severity in children with arthritis (Schanberg, Keefe, Lefebvre, Kredich, & Gill, 1998; Scott, Ansell, & Huskinsson, 1977), and postoperative recovery (Beyer & Aradine, 1987). VAS scores have been shown to be sensitive to changes in pain following analgesic medications (Romsing, Moller-Sonnergaard, Hertel, & Rasmussen, 1996) and psychological pain

management intervention (Powers, Blount, Bachanas, Cotter, & Swan, 1993).

Stability of any self-report pain rating is necessarily limited by the stability of the children's experience of pain. Therefore, high test–retest correlations would not necessarily be expected. For example, Gragg et al. (1996) obtained 6-month test–retest correlations of .41 for 8- to 16-year-old patients with rheumatic disease. Van Dijk et al. (2000) reported test–retest reliabilities of .58 in a sample of preschoolers tested at 3-hr intervals following surgery.

Advantages of VAS include ease of administration, low cost, and the fact that the scale yields ratio data. The VAS has been recommended as most appropriate for children over 8 years of age (Stinson et al., 2006). It is important to note that alterations in the VAS format can create significant changes in the psychometric properties of the scale. For example, Tesler et al. (1991) added the verbal labels of: ''little pain,'' ''medium pain,'' and ''large pain'' at regular points along a line anchored ''no pain'' and ''worst possible pain.'' The resulting scale, called the Word-Graphic Rating Scale, is an ordinal scale rather than a ratio scale; the distances between the verbal descriptions are not equal distance (Sinkin-Feldman, Tesler, & Savedra, 1997). Similar limitations apply to the Word Descriptor Scale (Whaley & Wong, 1987), which includes descriptors of: ''no pain,'' ''little,'' ''medium,'' ''quite a lot,'' ''very bad,'' and ''worst pain'' along a horizontal line.

Some authors have added numerical markers along the VAS (e.g., at 5- or 10-point intervals) (Jay, Elliott, Katz, & Siegel, 1987); however, this practice has been criticized because preferred or familiar numbers may be chosen more frequently (Huskisson, 1983; McGrath, 1990). Polkki et al. (2003) found that younger children understood the idea of matching a higher point with greater pain when using a vertically oriented VAS better than the left to right matching associated with a horizontal line VAS. However, Huskisson reported that individuals are more likely to select the extreme endpoints when the VAS is presented vertically rather than horizontally. VASs qualify for a rating of ''well-established.''

Faces scales consist of a set of line drawings or photographs of faces that depict pain states (Chambers et al., 1999). They are argued to be appropriate for children because they do not employ sophisticated words or abstract numerical values. Psychometrics have generally been strong for these scales. As examples, faces scales have been shown to correlate highly with other self-report indices (Spafford, von Baeyer, & Hicks, 2002), ratings by parents and nurses (Chambers et al., 1999), behavioral measures of pain (MacLaren & Cohen, 2005), and to be sensitive to analgesic and nonpharmacological interventions (Gold, Hyeon Kim, Kant, Joseph, & Rizzo, 2006; Spafford et al., 2002), to decrease during postprocedure (Paik & Ahn, 2002; Smith, Shah, Goldman, & Taddio, 2007), and to vary depending on the pain stimulus (e.g., different immunizations; Wood et al., 2004).

The affect portrayed in the anchors of the different faces scales is not consistent across scales and has been found to influence children's responses. For example, the smiling face anchor on the lower end of the Wong-Baker Faces Pain Rating Scale may bias children's pain scores, given that many children in medical situations are not smiling and happy (Chambers & Craig, 1998; Chambers, Hadial, Craig, Court, & Mongomery, 2005). On the other hand, the fact that the most distressed face of the six faces is crying may make some children not choose this face unless they themselves are crying (McCaffery, 2002).

McGrath (1987) also raised the concern that faces with numerical scales alongside them may confound the affective component of pain (presumably captured in the faces) and pain intensity (reflected in the numbers). Finally, it cannot be assumed that faces scales yield interval or ratio data, even if each face is labeled with an integer. At best, the data should be considered ordinal; the intervals between stimuli might well be unequal from the child's perspective (McGrath, de Veber, & Hearn, 1985).

A popular faces scale, *The Oucher* (Beyer, 1984, 2000; Beyer & Knott, 1998), provides different photographs for Caucasian, Hispanic, and Black children. The faces depicted in the *Wong-Baker Faces Scale* (Wong & Baker, 1988) and the *Faces Pain Scale-Revised* (FPS-R; Hicks et al., 2001), two other commonly used faces scale measures, are line drawings with no ethnicity distinctions. The Oucher and the Wong-Baker Scales assign a numerical value to each face (e.g., 0–5). The Wong-Baker Scale also adds word descriptors to each face (*no hurt*, *hurts a little*, *hurts a whole lot*, etc.) and the Oucher is oriented in a vertical orientation similar to a thermometer. The FPS-R contains six faces, ranging from a neutral expression to one of intense pain but without tears, and these faces can be numbered 0, 2, 4, 6, 8, and 10, which provides an approximation of the commonly used 0–10 metric. In addition, the FPS-R has been translated into over 30 languages and has received considerable empirical support. This group rated the Wong-Baker as ''approaching well-established,'' the

Oucher as ''well-established,'' and the Faces Pain Scale-Revised as ''well-established.''

The Poker Chip Tool or Pieces of Hurt Tool (Hester, 1979; Hester, Foster, & Kristensen, 1990) is a very simple pain intensity measure, consisting of four red chips that represent ''pieces of hurt'' (with one chip indicating a *little hurt* and all four chips indicating *the most hurt a child could have*). This straight-forward, concrete measure has been used with preschoolers from a variety of cultures, including Taiwanese (Suraseranivongse et al., 2005) and Jordanian (Gharaibeh & Abu-Saad, 2002) children. Scores tend to correlate with other self-report pain measures, such as the Oucher, word descriptors, and faces (Gharaibeh & Abu-Saad, 2002; Goodenough et al., 1997), and with behavioral observations of verbal, vocal, facial, and motor behaviors indicating pain (Hester). As would be expected, parent–child or nurse–child agreement using the poker chip tool typically is significant but moderate ($rs = .23 - .70$). Poker Chip Tool scores show the expected decreases following administration of analgesics (Romsing et al., 1996). The Poker Chip Tool appears to have the most utility as a simple clinical assessment tool to identify presence/absence of pain and very gross estimates of pain intensity in young children. It was recommended for use in 3- to 4-year-olds for acute procedure-related and postoperative pain (Stinson et al., 2006). The Poker Chip Tool was rated as ''well-established.''

Questionnaire and Diaries

Four self-monitoring measures were selected based on their frequency of use in the pain assessment literature. These self-monitoring measures are similar in terms of the informants completing them and the dimensions of pain assessed. However, the measures differ with respect to standardization sample or validation sample, psychometric properties, and the degree of cultural sensitivity. These measures are most typically used in populations of children with recurring or chronic pain that requires prospective monitoring.

The Headache Diary (Richardson et al., 1983) is a self-report measure using a 0–5 scale with instructions for the patient to complete ratings at specific times of the day (e.g., breakfast, lunch, dinner, and bedtime). The Headache Diary originated in Canada (Richardson et al., 1983) and has been used in the United States (Andrasik, Burke, Attanasio, & Rosenblum, 1985) and the Netherlands (Langeveld, Koot, & Passchier, 1999). The Headache Diary offers prospective ratings of pain and is easy to complete. However, diaries may contain fewer recordings than ideal due to noncompliance in completing the ratings (van de Brink, Bandell-Hoekstra, & Abu-Saad, 2001). In addition, discrepancies in ratings seem to occur for ''slight'' and ''mild'' ratings of pain, especially when changes in activity are stipulated (Richardson et al., 1983). The Headache Diary has shown sensitivity to treatment effects in headache pain intervention research (McGrath et al., 1992; Richter et al., 1986). The Headache Diary was rated as ''approaching well-established.''

The Pain Diary (Hunfeld et al., 2001; Hunfeld et al., 2002) employs a 100 mm horizontal line to assess the intensity, frequency, and duration of pain. It can be used across pain types. Similar to the Headache Diary, ratings are done at specified times of the day. Significant concordance between youths' and parents' ratings of headache activity on the Pain Diary has been reported for a 4-week period (Richardson et al., 1983) and following a 10-session intervention using either relaxation or temperature biofeedback (Andrasik et al., 1985). Whereas the Pain Diary may generate information about situations that might be associated with pain, it has focused on recurrent pain and only has a parent proxy form for children younger than 12 years of age (Hunfeld et al., 2001, 2002). The Pain Diary was assigned a rating of ''approaching well-established.''

The Abu-Saad Pediatric Pain Assessment Tool (PPAT, Abu-Saad et al., 1990) is a questionnaire designed to assess pain in school-age children. It was originally developed in the Netherlands and assesses multiple aspects of pain, such as triggers of pain and medication type and amount. It consists of 30 word descriptors of sensory, affective, and evaluative aspects of pain based on the work of Melzack and Torgerson (1971) and uses a 10 cm scale to measures present and worst pain. The PPAT has been administered to children with disease-related pain as well as to children undergoing surgery (Abu-Saad, 1994; Abu-Saad et al., 1990). Supporting the construct validity and clinical utility of the PPAT is evidence of changes in pain scores from pre- to post-analgesia administration (Abu-Saad, Pool, & Tulkens, 1994). In The PPAT has been administered to Arab-American (Abu-Saad, 1984) and Jordanian (Gharaibeh & Abu-Saad, 2002) children to determine cultural-specific aspects of pain assessment. In terms of reliability, the internal consistency of the PPAT (Abu-Saad et al., 1990) has been adequate. Significant correlations between child, parent, nurse, and physician ratings of present pain and worst pain have been found for the PPAT (Abu-Saad, 1994) and correlations between pain intensity ratings for the PPAT have

been variable and dependent on the patient population (i.e., juvenile rheumatoid arthritis, cancer, postoperative). However, correlations have not been significant between headache frequency and intensity on the Headache Diary and the 10 cm scale on the PPAT (van de Brink et al., 2001), as well as interview or questionnaire information (Laurell, Larson, & Eeg-Ologsson, 2003). The PPAT was assigned a rating of ''approaching well-established.''

The Varni/Thompson Pediatric Pain Questionnaire (PPQ; Varni & Thompson, 1985) is a multidimensional questionnaire for assessing childhood pain, with separate forms for the pediatric patient, parent, and clinician. It was modeled after the most widely used adult pain assessment instrument, the McGill Pain Questionnaire (Melzack, 1975). The PPQ assesses physician, patient, and parent perceptions of the patient's pain experience in a developmentally appropriate format. More specifically, this instrument measures pain intensity, location, and the sensory, evaluative, and affective qualities of the pain. The different forms use different formats, such as using colors for younger children and descriptive terms for adolescents. The clinical utility of the PPQ has primarily focused on two patient populations, JRA (Thompson, Varni, & Hanson, 1987) and sickle cell disease (SCD) (Walco & Dampier, 1990), and has been investigated in children with chronic pain (Walco, Sterling, Conte, & Engel, 1999). The PPQ has been translated into numerous languages, including Danish, French for France and for Canada, Norwegian, Portuguese for Brazil, Spanish for the United States, and Swedish (www.pedsql.org). In terms of psychometrics, good test–retest estimates for 1-week, 3-week, and 6-month intervals and correlations between child, parent, nurse, and physician ratings of present pain and worst pain have been found for the PPQ (Gragg et al.; Thompson et al.; Walco & Dampier). Furthermore, significant cross-informant ratings have been obtained but correlations have been higher between parent–child versus those with physicians. Lastly, for children with SCD, parent and physician VAS ratings were predictive of physician estimates of disability (Walco & Dampier, 1990). The workgroup rated the PPQ as ''well-established.''

Behavior Observational Scales for Assessing Pain and Behavioral Distress

Behavioral observation scales are used to assess patients' display of behaviors that are indicative of pain and distress. They utilize objective monitoring of operationally defined, subjectively displayed, and observable behaviors. There is an assumption that anxiety, fear, and other forms of behavioral distress are mixed with and co-existent with the experience of pain. Frequently, the terms behavioral distress and pain are used interchangeably. Indeed, the interwoven nature of the sensory aspects of pain experience and concomitant affective states is consistent with the definition of pain proposed by the International Association for the Study of Pain (IASP; 1979) as, ''Pain is an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such'' (p. 249).

Behavioral observation measures of pain are particularly applicable for younger children; children with cognitive impairments; or during times when the child is too distressed to indicate their level of pain, such as before, during, or shortly after painful injections (von Baeyer & Spagrud, 2007). Like self-report, behavioral manifestations of pain behaviors are also subject to the effects of minimization, exaggeration, and the influence of social and other contextual variables. Given the strengths and flaws inherent in any assessment method, we recommend that when appropriate, using a combination of behavioral observation and self-report methods.

With observational scales for adolescents, children, and infants, behaviors such as saying ''ouch,'' crying, the need for restraint, facial grimaces, and posturing are examples of socially agreed upon, empirically supported, and behavioral indices that reflect the experience of pain during medical procedures. Many of the behavioral observation scales were developed for studying pain and behavioral distress in children and adolescents with cancer who were experiencing painful bone marrow aspirations and lumbar puncture procedures. As such, they were designed to capture reactions to acute, intense, painful medical procedures that were often performed in a strong emotional and frequently social context (e.g., in the presence of medical staff and perhaps parents). The medical procedures had a defined start point, and observations occurred at specified times prior to, during, and after the medical procedure. Trained coders observe children's behavioral reactions or code videotapes of children's reactions to the procedures.

The Procedural Behavioral Rating Scale (PBRS) was developed by Katz, Kellerman, and Siegel in 1980. The PBRS originally included 25 behaviors, then 13, and eventually 11 in the revised version (PBRS-R; Katz, Kellerman, & Ellenberg, 1987). These 11 behaviors are *cry*, *cling*, *pain verbal*, *scream*, *stall*, *flail*, *refusal positions*,

restrain, *muscular rigidity*, *emotional support*, and *requests termination*. Each behavior was monitored as present or absent in each of four phases of the medical procedure. These behaviors were then totaled to obtain a PBRS total score. The disadvantage of the PBRS as originally designed is that it did not account for repeated displays of a behavior in the same phase, which would likely affect sensitivity. The PBRS has been very valuable as one of the first observational measures of distress. The PBRS was classified as ''approaching well-established.''

The Observational Scale of Behavioral Distress (OSBD, Jay et al., 1983) and the revised version (OSBD-R, Elliott et al., 1987) are among the most widely used scales in this category. The original OSBD contained 11 operationally defined behaviors that were indicative of distress. The number of distress behaviors on the OSBD-R was reduced to eight. Both the OSBD and OSBD-R are unique in that distress behaviors are weighted on a 1–4 point scale to reflect the intensity of distress. Examples include *information seeking* and *cry*, weighted at 1.5, and *scream*, *physical restraint*, and *flail*, weighted at 4. 0. Elliott et al. found higher correlations between nurses' and children's ratings of fear and pain when using weighted rather than unweighted scores. Behaviors are recorded as occurring or not occurring in 15 sec intervals within four phases of the medical procedure. Typically, these phase scores are added to produce a whole session distress score. The OSBD/OSBD-R was rated as ''well-established.''

The Child–Adult Medical Procedure Interaction Scale (CAMPIS; Blount et al., 1989) and the CAMPIS-Revised were developed in part based on the OSBD. The CAMPIS includes a child distress subscale on which behaviors are recorded based on their frequency of occurrence in phases that occur before, during, and after medical events. These behaviors may then be summed to obtain a total score per phase or grouped into categories, including apprehensive or anticipatory distress (*request emotional support*, *information seeking*, and *verbal fear*) and demonstrative distress (*cry*, *scream*, *verbal resistance*, *verbal pain*, and *verbal emotion*) (Blount, Sturges, & Powers, 1990). Additional behavior codes on the CAMPIS are indicative of child coping and other behaviors that children display during medical procedures. The CAMPIS also includes 17 behaviors that can be performed by parents, medical staff, or other adults who may accompany the child during medical treatments. The revised version of the CAMPIS groups these behaviors into six codes (CAMPIS-R; Blount et al., 1997, 1990). In addition to a *child distress* subscale, the other five CAMPIS-R codes are *child coping*, *child neutral*, *adult distress promoting*, *adult coping promoting*, and *adult neutral* behaviors. For measuring distress, both rates and proportions of the distress behaviors have been used, with good validity for both metrics (Blount et al., 1997). The CAMPIS-Short Form (Blount, Bunke, Cohen, & Forbes, 2001) was developed to provide an easier and quicker method of rating the CAMPIS-R codes. Recently, the CAMPIS has been modified for use in perioperative environments, with the introduction of the Perioperative Child–Adult Medical Procedure Interaction Scale (P-CAMPIS; Caldwell-Andrews, Blount, Mayes, & Kain, 2005) and for rating infant distress (CAMPIS-Infant Version, or CAMPIS-IV; Devine, Blount, Cheng, Seri, & Simons, 2006). Both P-CAMPIS and the CAMPIS-IV measure the same dimensions as the CAMPIS-R. The CAMPIS/CAMPIS-R was classified as ''well-established.''

The Procedure Behavior Checklist (PBCL; LeBaron & Zeltzer, 1984) is both a direct observation scale and a rating scale of behavioral distress. Eight behaviors are coded as occurring or not occurring before, during, and after the medical procedure. Occurrences are rated on a 1–5 point scale for intensity, with anchors of very mild to extremely intense. After the completion of the observation period, these ratings are added to yield a total PBCL score. The PBCL was classified as ''well-established.''

To avoid redundancy in reporting and given the consistency in findings across these measures, psychometrics for the PBRS, OSBD, CAMPIS, and PBCL will be described together subsequently. Across these four measures, inter-rater reliability is adequate to excellent (Blount et al., 1997; Chen, Craske, Katz, Schwartz, & Zeltzer, 2000; Jay et al., 1983; Katz et al., 1987) and internal consistency is good for the OSBD-R (Elliott et al., 1987). There is strong evidence supporting the validity of each of these measures. Distress scores have been found to correlate negatively with the age of the child (Jay et al.; LeBaron & Zeltzer, 1984), positively with anxiety (e.g., Jay et al.; Frank, Blount, Smith, Manimala, & Martin, 1995; LeBaron & Zeltzer), and positively with parental reports of the child's pain and distress during prior medical procedures (Jay et al.; Frank et al.). Children's distress during nonpainful anticipatory phases of the medical procedure predicted up to 86% of the variance of child distress during the painful phase of the BMA/LP procedures (Blount et al., 1990). Parental and medical staff behavior during the medical procedure has been found to account for 53% of the variance in child distress (Frank et al., 1995). Clinical utility has been shown by sensitivity to the effects of psychologically based and/or pharmacologically based interventions

(Cohen, Blount, Cohen, Schaen, & Zaff, 1999; Jay, Elliott, Fitzgibbons, Woody, & Siegel, 1995). Further, findings using additional subscales from the CAMPIS/CAMPIS-R have yielded results that directly inform the design of treatment intervention programs (Blount, Bunke, & Zaff, 2000).

The PBRS, OSBD, CAMPIS, and PBCL have each been used with pediatric patients undergoing various medical procedures, attesting to their broad applicability. These procedures include children undergoing laceration repair (Luhmann, Kennedy, Porter, Miller, & Jaffee, 2001), routine immunization injections (Blount et al., 1997), venipuncture (MacLaren & Cohen, 2005; Smith et al., 2007), burn injuries (Elliott & Olson, 1983), painful physical therapy regimens (Miller, Johanna-Murphy, & Zhelezniak, 2001), voiding cystourethogram procedures (Salmon & Pereira, 2002; Zelikovsky, Rodrigue, Gidyzc, & Davis, 2000), and outpatient surgery (Caldwell-Andrews et al., 2005). Further, parts of the CAMPIS-R have been used in analogue pain induction procedures with children, including the cold pressor task (Chambers, Craig, & Bennett, 2002) and a water loading procedure, in which children are instructed to drink as much water as they can until they are full (Walker et al., 2006). These observational measures also show applicability across ages (Jay et al., 1983), gender (Blount et al., 1997), languages (Calamenddrei et al., 1995), countries (Lossi & Hatira, 2003; Salmon & Pereira), and ethnicities including European-American (Blount et al., 1997); African-American (Cohen et al., 1999) and Latino-American (LeBaron & Zeltzer, 1984) samples.

In addition to the numerous strengths noted earlier, the main weakness of observational measures of behavioral distress is the time and effort required to complete them. For direct observation, coders must be trained and present during the time the behaviors are occurring or they must have access to video recordings of the children undergoing medical procedures. Although the information derived from them about the topography of distress is rich in detail, the amount of effort and personnel required may preclude ongoing usage in applied clinical settings. Rating scales that include some of the same dimensions, such as the PBCL and the CAMPIS-SF, may help to overcome the barrier to usage in clinical settings. However, recordings must be done with a trained observer.

In addition to the observational measures reviewed earlier, direct observation methods have proven to be particularly useful for assessing pain in infants and young children. They share many of the same strengths and weaknesses of the measures reviewed above.

The Children's Hospital of Eastern Ontario Pain Scale (CHEOPS; McGrath et al., 1985) is a measure of postoperative pain in children. It is distinct in that it can be used as a real-time, live-coding tool. It includes operational definitions for six domains: *cry*, *facial*, *child verbal*, *torso*, *touch*, and *legs*. Each domain is scored on a 0–3 scale. Inter-rater reliability following surgical procedures for children ages 1–5 years ranges from 90% to 99.2% for the different behavioral domains (McGrath et al.). CHEOPS scores correlate with child self-report of pain during injections, and it has been found to be sensitive to the effects of pharmacological interventions to reduce pain (Cassidy et al., 2001). However, no differences were detected on the CHEOPS between children receiving behavioral intervention and those who did not during immunizations (Cassidy et al., 2002). The CHEOPS was rated as ''well-established.''

The COMFORT Scale (Ambuel et al., 1992) was designed for use with children ages 0–18 years, although it has most often been used with children ages 0–5 years, who are in a pediatric intensive care unit. There are eight behaviorally anchored domains: *alertness*, *calmness/agitation*, *respiratory response*, *physical movement*, *mean arterial blood pressure*, *heart rate*, *muscle tone*, and *facial tension*. Each domain is rated using a 1–5 scale. Inter-rater reliability for the COMFORT total score using Pearson correlations was .84 (Ambuel et al.,). Two components of the COMFORT Scale have been identified, behavioral and physiological domains, which are strongly correlated. Validity is supported by correlations between COMFORT scores and VAS pain ratings (van Dijk et al., 2000). The COMFORT also was sensitive to differences displayed by 1-to-3-year-old children receiving either intermittent or continuous morphine (Bouwmeester et al., 2001) and to pre/postmorphine administration in babies weighing over 2,500 grams (Blauer & Gerstmann, 1998). The COMFORT Scale received a rating of ''well-established.''

The Child Facial Coding System (CFCS; Chambers et al., 1996) measures children's facial expressions during painful experiences. It was designed for preschool children, ages 2–5 years. It includes 13 facial actions (e.g., *eye squeeze*, *nasolabial furrow*), 10 of which are coded for intensity using the ratings of no action, slight action, and distinct/maximal action. Three items (i.e., *blink*, *flared nostril*, and *open lips*) are coded as present or absent. Inter-rater reliability ranged from .75 to .83 (Cassidy et al., 2002, Gilbert et al., 1999). CFCS scores differed depending on whether a patient was receiving an injection (Breau et al., 2001). The CFCS was

sensitive to group differences in pain expression for children receiving injections with or without EMLA (Cassidy et al., 2001) and for those receiving behavioral intervention or not (Cassidy et al., 2002). The CFCS has been used mostly in Canada and mostly with healthy children. Due to the detailed coding system, videotapes must be used to capture child facial actions. The CFCS was rated as ''approaching well established.''

The Premature Infant Pain Profile (PIPP; Stevens et al., 1996) is a 7-item observational, multidimensional measure of acute pain in preterm infants (gestational age between 28 and 36 weeks). Each item is scored on a 4-point scale. The items include physiological (heart rate, oxygen saturation) and behavioral dimensions (facial expression, eye squeeze, brow bulge, nasolabial furrow, and crying). Scores are summed across the seven items. Inter-rater reliability was high, with αs near .95 (Ballantyne, Stevens, McAllister, Dionne, & Jack, 1999). Scores on the PIPP for infants who experienced heel sticks were higher than for those who were held (Stevens et al.) or experienced a diaper change (Ballantyne et al.). The PIPP has been shown to be sensitive to various infant pain reducing interventions (Johnston, Stremler, Horton, & Friedman, 1999). The PIPP received a rating of ''well-established.''

The CHEOPS, COMFORT, CFCS, and PIPP have generally been shown to be sensitive to treatment effects, although to varying degrees. Due to their ease of use, the CHEOPS, COMFORT, and PIPP are applicable for real-time coding in clinical settings, and may help guide the application of pharmacological or other interventions.

Conclusions and Future Directions

Pediatric psychologists have made significant contributions to the development and validation of pain assessment tools, and are also frequent users of these instruments in their clinical and research practice. To date, there are a number of well-established assessment tools that can be used to evaluate pain in children of varying ages and with different types of pain in a variety of situations. However, the best measure will depend on the purpose of its use, the questions being asked, and the context in which it will be implemented. In this review, we provide an approach to evaluate measures to assess pain in infants, children, and adolescents, with an emphasis on examining the research-base and highlighting the relevant patient/situation context that should be considered when taking an evidence-based practice perspective.

As highlighted in the reviews, we wish to stress that the consumer of these and other reviews carefully consider the purpose of the tool. For example, if the measure is to be used in research or clinical work, different considerations will be in order. Furthermore, the specific type of research or clinical work and the populations be studied will also result in different priorities (feasibility, cost, psychometrics, and readability) when selecting an instrument. Given the increased focus on pediatric pain by larger medical governing bodies (e.g., JCAHO), we hope that the current review can guide healthcare professionals in the implementation of effective pain measures within their respective institutions.

There are several areas identified in the present review that require future investigation. For example, although considerable work has been done in devising developmentally appropriate rating scales for assessing pain intensity in children across the age span, additional efforts are necessary that take into account developmental and maturational factors beyond simply reading level. Further, much less attention has been devoted to validating measures of prospective monitoring of pain and functioning in children with recurrent and chronic pain. Novel methodologies have been emerging to enhance compliance in maintaining prospective records including the use of electronic pain diaries with children (Palermo, Valenzuela, & Stork, 2004; Stinson et al., 2006) that require future study. In addition, there is increasing research on the development of pain assessment tools for understanding pain expression and experience in cognitively impaired children (Stallard et al., 2002), and this remains an important area of measurement development and validation. Although measurement of functional disability, emotional functioning, and quality of life are recommended for understanding the specific impact of pain on children, these outcomes have not always been measured in response to intervention. A recent review of measures that capture the impact of chronic pain on adolescents was recently published (Eccleston, Jordan, & Crombez, 2006), highlighting the limited scope of measures available for pain impact on functioning (in contrast to pain symptoms).

In addition, pain assessment is limited regarding racial and ethnic differences. Research using large and representative samples would allow researchers to develop measures sufficiently sensitive to detect differences, if they exist. Lastly, assessment tools might quantify positive outcomes of pain, such as children's sense of self-efficacy following a brief stressor or enhanced pain coping or psychosocial growth that might develop along with

having chronic pain. Clearly, assessment development is needed to evaluate some of the rich contextual aspects of children's pain, which is critical in evidence-based assessment of pediatric pain.

Conflicts of interest: None declared.

Received March 9, 2007; revisions received September 16, 2007; accepted October 1, 2007

References

Abu-Saad, H. (1984). Cultural components of pain: The Arab-American child. *Issues in Comprehensive Pediatric Nursing, 7*, 91–99.

Abu-Saad, H. H. (1994). Pain in children: Developing a programme of research. *Disability and Rehabilitation, 16*, 45–50.

Abu-Saad, H. H., Kroonen, E., & Halfens, R. (1990). On the development of a multidimensional Dutch pain assessment tool for children. *Pain, 43*, 249–256.

Abu-Saad, H. H., Pool, H., & Tulkens, B. (1994). Further validity testing of the Abu-Saad Pediatric Pain Assessment Tool. *Journal of Advanced Nursing, 19*, 1063–1071.

Ambuel, B., Hamlett, K. W., Marx, C. M., & Blumer, J. L. (1992). Assessing distress in pediatric intensive care environments: The COMFORT scale. *Journal of Pediatric Psychology, 17*, 95–109.

Andrasik, F., Burke, E. J., Attanasio, V., & Rosenblum, E. L. (1985). Child, parent, and physician reports of a child's headache pain: Relationships prior to and following treatment. *Headache, 25*, 421–425.

APA presidential task force on evidence-based practice (2006). Evidence-based practice in psychology. *American Psychologist, 61*, 271–285.

Aradine, C. R., Beyer, J. E., & Thompson, J. M. (1988). Children's pain perception before and after analgesia: A study of instrument construct validity and related issues. *Journal of Pediatric Nursing, 3*, 11–23.

Ballantyne, M., Stevens, B., McAllister, M., Dionne, K., & Jack, A. (1999). Validation of the premature infant pain profile in the clinical setting. *The Clinical Journal of Pain, 15*, 297–303.

Beyer, J. E. (1984). *The Oucher: A user's manual and technical report*. Evanston, IL: The Hospital Play Equipment Company.

Beyer, J. E. (2000). Judging the effectiveness of analgesia for children and adolescents during vaso-occlusive events of sickle cell disease. *Journal of Pain and Symptom Management, 19*, 63–72.

Beyer, J. E., & Aradine, C. R. (1987). Patterns of pediatric pain intensity: A methodological investigation of a self-report scale. *The Clinical Journal of Pain, 3*, 130–141.

Beyer, J. E., & Knott, C. B. (1998). Construct validity estimation for the African-American and Hispanic versions of the Oucher Scale. *Journal of Pediatric Nursing, 13*, 20–31.

Blauer, T., & Gerstmann, D. (1998). A simultaneous comparison of three neonatal pain scaled during common NICU procedures. *The Clinical Journal of Pain, 14*, 39–47.

Blount, R. L., Bunke, V. L., Cohen, L. L., & Forbes, C. J. (2001). The Child-Adult Medical Procedure Interaction Scale-Short Form (CAMPIS-SF): Validation of a rating scale for children's and adults' behaviors during painful medical procedures. *Journal of Pain and Symptom Management, 22*, 591–599.

Blount, R. L., Bunke, V. L., & Zaff, J. F. (2000). Bridging the gap between explicative and treatment research: A model and practical implications. *Journal of Clinical Psychology in Medical Settings, 7*, 79–90.

Blount, R. L., Cohen, L. L., Frank, N. C., Bachanas, P. J., Smith, A. J., Manimala, M. R., et al. (1997). The Child-Adult Medical Procedure Interaction Scale-Revised: An assessment of validity. *Journal of Pediatric Psychology, 22*, 73–88.

Blount, R. L., Corbin, S. M., Sturges, J. W., Wolfe, V. V., Prater, J. M., & James, L. D. (1989). The relationship between adults behavior and child coping and distress during BMA/LP procedures: A sequential analysis. *Behavior Therapy, 20*, 585–601.

Blount, R. L., Sturges, J. W., & Powers, S. W. (1990). Analysis of child and adult behavioral variations by phase of medical procedure. *Behavior Therapy, 21*, 33–48.

Bouwmeester, N. J., Anand, K. J., van Dijk, M., Hop, W. C., Boomsma, F., & Tibboel, D. (2001). Hormonal and metabolic stress responses after major surgery in children aged 0–3 years: A double-blind, randomized trial comparing the effects of continuous versus intermittent morphine. *British Journal of Anaesthesia, 87*, 390–399.

Breau, L. M., McGrath, P. J., Craig, K. D., Santor, D., Cassidy, K. L., & Reid, G. J. (2001). Facial

expression of children receiving immunizations: A principal components analysis on the child facial coding system. *The Clinical Journal of Pain, 17*, 178–186.

Calamendrei, M., Messeri, A., Busoni, P., Bernini, G., Lippi, A., & Tucci, F. (1995). Comparison of two application techniques of EMLA and pain assessment in pediatric oncology patients. *Regional Anesthesia, 21*, 557–560.

Caldwell-Andrews, A. A., Blount, R. L., Mayes, L. C., & Kain, Z. N. (2005). Behavioral interactions in the perioperative environment: A new conceptual framework an the development of the Perioperative Child-Adult Medical procedure Interaction Scale. *Anesthesiology, 103*, 1130–1135.

Cassidy, K. L., Reid, G. J., McGrath, P. J., Finley, G. A., Smith, D. J., Morley, C., et al. (2002). Watch needle, watch TV: Audiovisual distraction in preschool immunization. *Pain Medicine, 3*, 108–118.

Cassidy, K. L., Reid, G. J., McGrath, P. J., Smith, D. J., Brown, T. L., & Finley, G. A. (2001). A randomized double-blind, placebo-controlled trial if the EMLA® patch for the reduction of pain associated with intramuscular injections in four to six-year-old children. *Acta Paediatrica, 90*, 1329–1336.

Chambers, C., & Craig, K. D. (1998). An intrusive impact of anchors in children's faces pain scales. *Pain, 78*, 23–37.

Chambers, C. T., Cassidy, K. L., McGrath, P. J., Gilbert, C. A., & Craig, K. D. (1996). *Child Facial Coding System Manual*. Halifax, Nova Scotia: Dalhousie University and University of British Columbia. Unpublished manuscript.

Chambers, C. T., Craig, K. D., & Bennett, S. M. (2002). The impact of maternal behavior on children's pain experience: An experimental analysis. *Journal of Pediatric Psychology, 27*, 293–301.

Chambers, C. T., Giesbrecht, K., Craig, K. D., Bennett, S. M., & Huntsman, E. (1999). A comparison of faces scales for the measurement of pediatric pain: Children's and parents' ratings. *Pain, 83*, 25–35.

Chambers, C., Hadial, J., Craig, K., Court, C., & Mongomery, C. (2005). Faces pain scales for the measurement of postoperative pain intensity in children following minor surgery. *Journal of Pain, 21*, 277–285.

Chambers, C. T., Reid, G. J., McGrath, P. J., & Finley, G. A. (1996a). Development and preliminary validation of a postoperative pain measure for parents. *Pain, 68*, 307–313.

Chen, E., Craske, M. G., Katz, E. R., Schwartz, E., & Zeltzer, L. K. (2000). Pain-sensitive temperament: Does it predict procedural distress and response to psychological treatment among children with cancer? *Journal of Pediatric Psychology, 25*, 269–278.

Cohen, L. L., Blount, R. L., Cohen, R. J., Schaen, E. R., & Zaff, J. F. (1999). Comparative study of distraction versus topical anesthesia for pediatric pain management during immunizations. *Health Psychology, 18*, 591–598.

Cohen, L. L., La Greca, A. M., Blount, R. L., Kazak, A. E., Holmbeck, G. N., & Lemanek, K. L. (May 4, 2007). Introduction: Evidence-based assessment in pediatric psychology. *Journal of Pediatric Psychology*. doi:10.1093/jpepsy/jsj115.

Cohen, L. L., MacLaren, J. E., & Lim, C. S. (2007). Pain and pain management. In R. G. Steele, T. D. Elkin, & M. C. Roberts (Eds.), *Handbook of evidence based therapies for children and adolescents*. New York: Springer Publishers, *In press*.

Devine, K., Blount, R., Cheng, P., Seri, L., & Simons, L. (2006). The impact of adult behaviors and vocalizations on infant distress during immunization. In R. L. Blount (Chair), *Adult influences on children's pain and distress: Expanding the paradigm*. Symposium conducted at the Seventh International Symposium on Pediatric Pain, Vancouver, CA.

Dworkin, R. H., Turk, D. C., Farrar, J. T., Haythornthwaite, J. A., Jensen, M. P., Katz, N. P., et al. (2005). Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain, 113*, 9–19.

Eccleston, C., Jordan, A. L, & Crombez, G. (2006). The impact of chronic pain on adolescents: A review of previously used measures. *Journal of Pediatric Psychology, 31*, 684–697.

Elliott, C. H., Jay, S. M., & Woody, P. (1987). An observation scale for measuring children's distress during medical procedures. *Journal of Pediatric Psychology, 12*, 543–551.

Elliott, C. H., & Olson, R. A. (1983). The management of children's distress in response to painful medical treatment for burn injuries. *Behaviour Research and Therapy, 21*(6), 675–683.

Finley, G. A., & McGrath, P. J. (1998). Introduction: The roles of measurement in pain management and research. In G. A. Finley, & P. J. McGrath (Eds.),

Measurement of pain in infants and children (pp. 1–4). Seattle, WA: IASP Press.

Frank, N. C., Blount, R. L., Smith, A. J., Manimala, R., & Martin, J. K. (1995). Parent and staff behavior, previous child medical experience, and maternal anxiety as they relate to child procedural distress and coping. *Journal of Pediatric Psychology, 20*, 277–289.

Gharaibeh, M., & Abu-Saad, H. (2002). Cultural validation of pediatric pain assessment tools: Jordanian perspective. *Journal of Transcultural Nursing, 13*, 12–18.

Gilbert, C. A., Lilley, C. M., Craig, K. D., McGrath, P. J., Court, C. A., Bennett, S. M., et al. (1999). Postoperative pain expression in preschool children: Validation of the child facial coding system. *Clinical Journal of Pain, 15*, 192–200.

Gold, J. I., Hyeon Kim, S. H., Kant, A. J., Joseph, M. H., & Rizzo, A. (2006). Effectiveness of virtual reality for pediatric pain distraction during IV placement. *CyberPsychology & Behavior, 9*(2), 207–212.

Goodenough, B., Addicoat, L., Champion, G. D., McInerney, M., Young, B., Juniper, K., et al. (1997). Pain in 4–6-year-old children receiving intramuscular injections: A comparison of the Faces Pain Scale with other self-report and behavioral measures. *The Clinical Journal of Pain, 13*, 60–73.

Gragg, R. A., Rapoff, M. A., Danovsky, M. B., Lindsley, C. B., Varni, J. W., Waldron, S. A., et al. (1996). Assessing chronic musculoskeletal pain associated with rheumatic disease: Further validation of the pediatric pain questionnaire. *Journal of Pediatric Psychology, 21*, 237–250.

Guariso, G., Mozrzymas, R., Gobber, D., Benini, F., Zancan, L., & Zacchello, F. (1990). Self-report assessment of recurrent abdominal pain. *Medical and Surgical Pediatrics, 21*, 255–260.

Hester, N. K. (1979). The preoperational child's reaction to immunization. *Nursing Research, 28*(4), 250–255.

Hester, N. O., Foster, R., & Kristensen, K. (1990). Measurement of pain in children: Generalizability and validity of the pain ladder and the poker chip tool. In D. C. Tyler, & E. J. Crane (Eds.), *Advances in pain research and therapy* (Vol. 15, pp. 79–84). New York: Raven.

Hicks, C. L., von Baeyer, C. L., Spafford, P., van Korlaar, I., & Goodenough, B. (2001). The Faces Pain Scale – revised: Toward a common metric in pediatric pain measurement. *Pain, 93*, 173–183.

Hunfeld, J. A. M., Perquin, C. W., Duivenvoorden, H. J., Hazebroek-Kampschreur, A. A. J. M., Passchier, J., van Suijlekom-Smit, L. W. A., et al. (2001). Chronic pain and its impact on quality of life in adolescents and their families. *Journal of Pediatric Psychology, 26*, 145–153.

Hunfeld, J. A. M., Perquin, C. W., Hazebroek-Kampschreur, A. A. J. M., Passchier, J., van Suijlekom-Smit, L. W. A., & van der Wouden, J. C. (2002). Physically unexplained chronic pain and its impact on children and their families: The mother's perception. *Psychology and Psychotherapy, 75*, 251–260.

Hunfeld, J. A. M., van der Wouden, J. C., den Deurwaarder, E. S. G., van Suijlekom-Smit, L. W. A. C. W., & Hazebroek-Kampschreur, A. A. J. M. (1997). Measuring chronic pain in children: An exploration. *Perceptual and Motor Skills, 84*, 1176–1178.

Huskisson, E. C. (1983). Visual analogue scales. In R. Melzack (Ed.), *Pain measurement and assessment* (pp. 33–37). New York NY: Raven Press.

IASP Subcommittee on Taxonomy (1979). Pain terms: A list with definitions and notes on usage: Recommended by the IASP Subcommittee on Taxonomy. *Pain, 6*, 249.

Jay, S., Elliott, C. H., Fitzgibbons, I., Woody, P., & Siegel, L. (1995). A comparative study of cognitive behavior therapy versus general anesthesia for painful medical procedures in children. *Pain, 62*, 3–9.

Jay, S. M., Elliott, C. H., Katz, E., & Siegel, S. E. (1987). Cognitive-behavioral and pharmacologic interventions for children's distress during painful medical procedures. *Jouranl of Consulting and Clinical Psychology, 55*, 860–865.

Jay, S. M., Ozolins, M., Elliott, C. H., & Caldwell, S. (1983). Assessment of children's distress during painful medical procedures. *Health Psychology, 2*, 133–147.

Johnston, C., Stremler, R., Horton, L., & Friedman, A. (1999). Effect of repeated doses of sucrose during heel stick procedure in preterm neonates. *Biology of the Neonate, 75*, 160–166.

Katz, E. R., Kellerman, J., & Ellenberg, L. (1987). Hypnosis in the reduction of acute pain and distress in children with cancer. *Journal of Pediatric Psychology, 12*(3), 379–394.

Katz, E. R., Kellerman, J., & Siegel, S. E. (1980). Behavioral distress in children with cancer undergoing medical procedures: Developmental

considerations. *Journal of Consulting and Clinical Psychology, 48*, 356–265.

Langeveld, J. H., Koot, H. M., & Passchier, J. (1999). Do experienced stress and trait negative affectivity moderate the relationship between headache and quality of life in adolescents? *Journal of Pediatric Psychology, 24*, 1–11.

Laurell, K., Larson, B., & Eeg-Ologsson, O. (2003). Headache in schoolchildren: Agreement between different sources of information. *Cephalalgia, 23*, 420–428.

LeBaron, S., & Zeltzer, L. (1984). Assessment of acute pain and anxiety in children and adolescents by self-reports, observer reports, and a behavior checklist. *Journal of Consulting and Clinical Psychology, 52*(5), 729–738.

Lossi, C., & Hatira, P. (2003). Clinical hypnosis in the alleviation of procedure-related pain in pediatric oncology patients. *The International Journal of Clinical & Experimental Hypnosis, 51*(1), 4–28.

Luffy, R., & Grove, S. K. (2003). Examining the validity, reliability, and preference of three pediatric pain measurement tools in African American children. *Pediatric Nursing, 29*, 54–59.

Luhmann, J. D., Kennedy, R. M., Porter, F. L., Miller, J. P., & Jaffee, D. M. (2001). A randomized clinical trial of continuous-flow nitrous oxide and midazolam for sedation of young children during laceration repair. *Annals of Emergency Medicine, 37*, 20–27.

MacLaren, J. E., & Cohen, L. L. (2005). A comparison of distraction strategies for venipuncture distress in children. *Journal of Pediatric Psychology, 30*, 387–396.

McCaffery, M. (2002). Choosing a faces pain scale. *Nursing, 32*, 68.

McGrath, P., de Veber, L., & Hearn, M. (1985). Multidimensional pain assessment in children. *Advances in Pain Research and Therapy, 9*, 387–393.

McGrath, P. A. (1987). An assessment of children's pain: A review of behavioral, physiological and direct scaling techniques. *Pain, 31*, 147–171.

McGrath, P. A. (1990). Pain assessment in children – a practical approach. In D. C. Tyler, & E. J. Krane (Eds.), *Advances in pain research therapy* (Vol. 15). New York: Raven Press, Ltd.

McGrath, P. J., Humphreys, P., Keene, D., Goodman, J. T., Lascelles, M. A., Cunningham, S. J., et al. (1992). The efficacy and efficiency of a self-administered treatment for adolescent migraine. *Pain, 49*, 321–324.

McGrath, P. J., Johnson, G., Goodman, J. T., Dunn, J., & Chapman, J. (1985). CHEOPS: A behavioral scale for rating postoperative pain in children. In H. L. Fields, R. Dubner, & F. Cervero (Eds.), *Advances in pain research and therapy* (Vol. 9, pp. 395–402). New York: Raven Press.

Melzack, R. (1975). The McGill pain questionnaire: Major properties and scoring methods. *Pain, 1*, 277–299.

Melzack, R., & Torgerson, W. S. (1971). On the language of pain. *Anesthesia, 34*, 50–59.

Melzack, R., & Wall, P. D. (1965). Pain mechanisms: A new theory. *Science, 150*, 971–979.

Merkel, S. I., Voepel-Lewis, T., Shayevitz, J. R., & Malviya, S. (1997). The FLACC: A behavioral scale for scoring postoperative pain in young children. *Pediatric Nursing, 23*, 293–297.

Miller, A. C., Johanna-Murphy, M., & Zhelezniak, V. (2001). Impact of the therapist-child dyad on children's pain and coping during medical procedures. *Developmental Medicine & Child Neurology, 43*, 118–123.

O'Rourke, D. (2004). The measurement of pin in infants, children, and adolescents: From policy to practice. *Physical Therapy, 84*, 560–570.

Paik, H. J., & Ahn, Y. M. (2002). Measurement of acute pain after eye surgery in children. *Korean Journal of Ophthalmology, 16*, 103–109.

Palermo, T. M., Valenzuela, D., & Stork, P. P. (2004). A randomized trial of electronic versus paper pain diaries in children: Impact on compliance, accuracy, and acceptability. *Pain, 107*, 213–219.

Phillips, D. M. (2000). JCAHO pain management standards are unveiled. *The Journal of the American Medical Association, 284*, 428–429.

Polkki, T., Pietila, A., & Vehvilainen-Julkunen, K. (2003). Hospitalized children's descriptions of their experiences with postsurgical pain relieving methods. *International Journal of Nursing Studies, 40*, 33–44.

Powers, S. W., Blount, R. L., Bachanas, P. J., Cotter, M. W., & Swan, S. C. (1993). Helping preschool leukemia patients and their parents cope during injections. *Journal of Pediatric Psychology, 18*, 681–695.

Richardson, G. M., McGrath, P. J., Cunningham, S. J., & Humphreys, P. (1983). Validity of the headache diary for children. *Headache, 23*, 184–187.

Richter, I. L., McGrath, P. J., Humphreys, P. J., Goodman, J. T., Firestone, P., & Kenne, D. (1986). Cognitive and relaxation treatment of paediatric migraine. *Pain, 25*, 195–203.

Romsing, J., Moller-Sonnergaard, J., Hertel, S., & Rasmussen, M. (1996). Postoperative pain in children: Comparison between ratings of children and nurses. *Journal of Pain and Symptom Management, 11*(1), 42–46.

Salmon, K., & Pereira, J. K. (2002). Predicting children's response to an invasive medical investigation: the influence of effortful control and parent behavior. *Journal of Pediatric Psychology, 27*, 227–233.

Schanberg, L., Keefe, F., Lefebvre, J., Kredich, D., & Gil, K. (1998). Social context of pain in children with juvenile primary fibromyalgia syndrome: parental pain history and family environment. *The Clinical Journal of Pain, 14*, 107–115.

Scott, P., Ansell, B., & Huskisson, E. (1977). Measurement of pain in juvenile chronic polyarthritis. *Annals of the Rheumatic Diseases, 36*, 186–187.

Sinkin-Feldman, L., Tesler, M., & Savedra, M (1997). Word placement on the Word-Graphic Rating Scale by pediatric patients. *Pediatric Nursing, 23*, 31–34.

Smith, R. W., Shah, V., Goldman, R. D., & Taddio, A. (2007). Caregivers' responses to pain in their children in the emergency department. *Archives of Pediatric and Adolescent Medicine, 161*, 578–582.

Spafford, P. A., von Baeyer, C. L., & Hicks, C. L. (2002). Expected and reported pain in children undergoing ear piercing: A randomized trial of preparation by parents. *Behaviour Research and Therapy, 40*, 253–266.

Stallard, P., Williams, L., Velleman, R., Lenton, S., McGrath, P. J., & Taylor, G. (2002). The development and evaluation of the pain indicator for communicatively impaired children (PICIC). *Pain, 98*, 145–149.

Stevens, B., Johnston, C., Petryshen, P., & Taddio, A. (1996). Premature infant pain profile: Development and initial validation. *Clinical Journal of Pain, 12*, 13–22.

Stewart, W. F., Ricci, J. A., Chee, E., Morganstein, D., & Lipton, R. (2003). Lost productive time and cost due to common pain conditions in the US workforce. *The Journal of the American Medical Association, 290*, 2443–2454.

Stinson, J. N., Kavanagh, T., Yamada, J., Gill, N., & Stevens, B. (2006). Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. *Pain, 125*, 143–157.

Stinson, J. N., Petroz, G. C., Tait, G., Feldman, B. M., Streiner, D., McGrath, P. J., et al. (2006). e-Ouch: Usability testing of an electronic chronic pain diary for adolescents with arthritis. *Clinical Journal of Pain, 22*, 295–305.

Suraseranivongse, S., Montapaneewat, T., Monon, J., Chainhop, P., Petcharatana, S., & Kraiprasit, K. (2005). Cross-validation of a self-report scale for postoperative pain in school-aged children. *Journal of the Medical Association of Thailand, 88*(3), 412–418.

Tesler, M. D., Savedra, M. C., Hlzener, W. L., Wilkie, D. J., Ward, J. A., & Paul, S. M. (1991). The Word-Graphic Rating Scale as a measure of children's and adolescents' pain intensity. *Research in Nursing and Health, 14*, 361–371.

Thompson, K. L., Varni, J. W., & Hanson, V. (1987). Comprehensive assessment of pain in juvenile rheumatoid arthritis: An empirical model. *Journal of Pediatric Psychology, 12*, 241–255.

van de Brink, M., Bandell-Hoekstra, E. N. G., & Abu-Saad, H. H. (2001). The occurrence of recall bias in pediatric headache: A comparison of questionnaire and diary data. *Headache, 41*, 11–20.

van Dijk, M., de Boer, J., Koot, H., Tibboel, D., Passchier, J., & Duivenvoorden, H. (2000). The reliability and validity of the COMFORT scale as a postoperative pain instrument in 0 to 3-year-old infants. *Pain, 84*, 367–377.

Varni, J. W., & Thompson, K. L. (1985). The Varni/Thompson Pediatric Pain Questionnaire. Unpublished manuscript.

Varni, J. W., Thompson, K. L., & Hanson, V. (1987). The Varni/Thompson Pediatric Pain Questionnaire: I. Chronic musculoskeletal pain in juvenile rheumatoid arthritis. *Pain, 28*, 27–38.

von Baeyer, C. L., & Spagrud, L. J. (2007). Systematic review of observational (behavioral) measures of pain for children and adolescents aged 3 to 18 years. *Pain, 127*, 140–150.

Walco, G. A., & Dampier, C. D. (1990). Pain in children and adolescents with sickle cell disease: A descriptive study. *Journal of Pediatric Psychology, 15*, 643–658.

Walco, G. A., Sterling, C. M., Conte, P. M., & Engel, R. G. (1999). Empirically supported treatments in pediatric psychology: Disease-related pain. *Journal of Pediatric Psychology, 24*, 155–163.

Walker, L. S., Williams, S. E., Smith, C. A., Garber, J., Van Slyke, D. A., & Lipani, T. A. (2006). Parent attention versus distraction: Impact on symptom complaints by children with and without functional abdominal pain. *Pain, 122*, 43–52.

Whaley, L. F., & Wong, D. (1987). *Nursing care of infants and children* (3rd ed.), St. Louis: Mosby Year Book.

Wong, D. L., & Baker, C. (1988). Pain in children: Comparison of assessment scales. *Pediatric Nursing, 14*, 9–17.

Wood, C., von Baeyer, C. L., Bourrillon, A., Dejos-Conant, V., Clyti, N., & Abitbol, V. (2004). Self-assessment of immediate post-vaccination pain after two different MMR vaccines administered as a second dose in 4- to 6-year-old children. *Vaccine, 23*, 127–131.

Young, K. D. (2005). Pediatric procedural pain. *Annals of Emergency Medicine, 45*, 160–171.

Zelikovsky, N., Rodrigue, J. R., Gidyzc, C. A., & Davis, M. A. (2000). Cognitive behavioral and behavioral interventions help young children cope during a voiding cystourethrogram. *Journal of Pediatric Psychology, 25*, 535–543.