

Note: This article was originally published in *The Suffering Child* 2005;8 (June) but that online journal has since been discontinued. The paper is now available at the following web address:  
<http://www.usask.ca/childpain/research/6scales/6scales.pdf>

## Comparing six self-report measures of pain intensity in children

Belinda Goodenough<sup>a,b</sup>, PhD

Tiina Piira<sup>b,c</sup>, MPsychol(Clinical)

Carl L. von Baeyer<sup>d</sup>, PhD

Karen Chua<sup>b</sup>, BSc(Psychol)

Eunice Wu<sup>c</sup>, BSc(Psychol)

Joseph D.H. Trieu<sup>c</sup>, MBBS

G. David Champion<sup>c</sup>, MD

a School of Women's and Children's Health, University of New South Wales, Sydney, Australia

b School of Psychology, University of New South Wales, Sydney, Australia

c Pain Research Unit, Sydney Children's Hospital, Randwick, Australia

d Departments of Psychology and Pediatrics, University of Saskatchewan, Saskatoon, Canada

### Corresponding author:

Tiina Piira

Pain Research Unit

Sydney Children's Hospital,

High Street

Randwick, NSW 2031

Australia

Tel: +61 2 9382 1585; Fax: +61 2 9382 1008

Email: PiiraT@sesahs.nsw.gov.au

**Key words:** children, measurement, pain, scale, self-report

**Running title:** Measures of pain intensity in children

### Abstract

Many tools for self-report measurement of children's pain have been developed independently of each other. There has been little research comparing their properties so that the best tools can be chosen. The present study compared and cross-calibrated six self-report measures of children's pain intensity. The measures examined were the Faces Pain Scale (FPS), the Coloured Analogue Scale (CAS), the Pieces of Hurt Tool, an Adjectival Scale, the Finger Span Measure, and the Sydney Animated Facial Expression (SAFE) scale. Participants were 82 inpatients, aged 4 to 16 years, who had experienced some pain during their hospital admission. They rated their pain severity using each measure and indicated the score at which they thought a nurse should be called for help. For analyses of mean differences across measures, each was standardized to a 0-to-10 metric. Scores on the six scales were quite highly inter-correlated (median  $r = .65$ ). However, there was greater variability in standardized scores across scales for younger rather than older children, suggesting that younger children were more influenced by specific scale characteristics or that they had less understanding of the scale and the self-report task. The SAFE was rated by children as the easiest to use, followed by the FPS and CAS. On the basis of internal consistency, response distribution, child preference, and cost, the FPS (or the Faces Pain Scale – Revised) and the CAS are recommended.

## Introduction

Children are often under-treated for pain, especially after surgery [10,14,21,26,31,37,40]. A significant contributing factor is likely to be the difficulty that health professionals may have in adequately assessing children's pain. As a partial response to this issue, a variety of methods have been developed to help measure pain intensity in children, taking into account the cognitive abilities of the child. These have included self-report, behavioral (observational), and physiological methods [8].

Despite the apparent objectivity of physiological measures, no single physiological index has been found to be a reliable measure of pain. Behavioral observation methods also have considerable limitations, as overt behavior may not reflect the subjective experience of pain, particularly in long-lasting pain. Therefore, clinicians have been exhorted to obtain and rely on self-report measures whenever possible [27], especially when evaluating the efficacy of treatments [28].

The experience of pain is complex. Potentially measurable qualities include sensory intensity, spatial location, temporal duration, affect and distress, as well as appraisal of the meaning of the pain and interference with quality of life. Pain intensity is the most commonly measured dimension, with a wide variety of self-report measures developed for children. These include scales based on numerical ratings, facial expression, verbal descriptors, and age-appropriate derivatives of visual analog scales [8].

This wide variety of pediatric measures reflects a proper concern for standardized age-appropriate assessment of pain, and application of relevant data from cognitive-developmental psychology in deriving and testing those measures. Yet the array of choices can be problematic. For example, treatment outcomes are difficult to compare

when based upon different methods and metrics. There is a risk of limited generalizability of research results or clinical practice guidelines beyond the institutional setting.

Furthermore, there is significant variability in the known psychometric properties of different pain measurement tools and in their scoring metrics [35]. These issues may confuse and undermine the confidence of clinicians and researchers in deciding which pain measure to use.

In the absence of recommendations for a single scale type as a standard for use with children, one approach is to generate data relevant to cross-calibration of very different pain scale formats. This approach is justified given research showing that even subtle variations in defining scale anchor points can influence pain ratings. For example, on facial expression scales, mean pain ratings by both children and parents may be significantly higher if the zero-pain face projects a smiling rather than a neutral expression [6]. There are also data suggesting that some children, at least in the 4- to 6-year age group, do not apply the anchor-points of 'most pain' and 'zero pain' consistently across different self-report scale formats [18].

Another under-researched aspect of cross-calibration of pain scales for children is the consideration of a clinically meaningful value relevant to intervention (e.g., administering analgesia). For self-report measures this scale value may correspond with the magnitude of pain for which a patient expects or wishes to receive analgesia [12]. This third type of anchor point is likely to be highly subjective, but is an important source of information for understanding how children may use scales of different types to communicate with care-givers.

The present study explored some of the psychometric properties and the clinical utility of six different self-report measures of pain intensity within a pediatric clinical context. Although any of a large number of possible measures could have been selected for use in this study, the scales were chosen to represent a number of different classes of self-report pain intensity measures, namely those based on facial expression, visual analog, concrete ordinal, and verbal rating scales. The specific measures being investigated were the Faces Pain Scale (FPS) [2], the Coloured Analogue Scale (CAS) [29], the Pieces of Hurt Tool [23], the Finger Span Measure [16], an Adjectival Scale, and the Sydney Animated Facial Expression Scale (SAFE) [19,24]. These tools represent highly diverse response formats for children: pointing to a face, moving a sliding marker, selecting pieces of hurt, making a folk psychology finger gesture, choosing a verbal descriptor, and pressing buttons on a computer keyboard.

The six measures were compared in several respects. First, the level of inter-scale concordance [33] across each of the pain intensity measures was assessed. Second, the scales were calibrated with a clinically meaningful indicator, namely the point (i.e., rating on the scale) at which the child thought a nurse should be called to help with the pain. Third, the response distributions for each scale were examined, particularly with regard to any possible end-point bias. To determine which measures were associated with relatively higher and lower ratings, average pain intensity ratings were also compared with scores on all measures standardized to a 0-to-10 scale. This method of standardized comparison has been used in other studies [34] and does not affect the correlations in any way [9]. Finally, practical considerations were explored as indexed by the child's rank-order appraisal of the six scales on ease of use. Within these comparisons, possible age

differences were examined, given that developmental differences may impact on children's ability to use various assessment tools and the tendency for younger children to display certain response biases [5,7,8].

## **Materials and Methods**

### *Participants*

Participants were 82 pediatric inpatients between the ages of 4 and 16 years (35 males and 47 females) at the Sydney Children's Hospital, Randwick, Australia. Approximately 100 consecutive inpatients from three general medical and surgical wards in the hospital were invited to participate. A broad inclusion criterion was that the child had undergone a procedure (surgical or medical) involving pain, or received assessment or treatment of a painful condition. Children were excluded if a parent was not present during the painful experience – additional data were also being collected from parents for a related project (not reported in the current paper). The mean age of children participating was 9.2 years ( $SD = 3.53$ ). The sample was divided into younger (4-7 year old,  $n = 37$ ) and older (8-16 year old,  $n = 45$ ) children to explore age-related differences. This age categorization was based on previous research suggesting that pain intensity ratings may not be a linear function of age but may plateau after about the age of 7-8 years [20].

The painful contexts on which children were reporting included medical procedures (35.3 %), post-operative pain (16.5 %), recurrent or persistent aches and pains including disease-related pains (41.2 %), and accidents or injuries (4.7 %). Children provided estimates of the maximum pain intensity experienced during that event or

episode. If the child was in any pain at the time the researcher visited, then scores for that present pain were requested; this was the case for 11 (13.6%) in the sample.

### *Materials*

Six pain intensity measures were used.

**The Faces Pain Scale** (FPS; *see footnote 1*) [2] is a measure that assesses pain intensity by asking the child to select which of the seven line-drawn faces best corresponds with his/her level of pain. The faces range from a neutral expression, corresponding to “no pain” on the far left (scored 0), to a face depicting “most pain possible” on the far right (scored 6). The faces depict physical features similar to real human pain expressions, such as brow furrow and horizontal mouth stretch. Physical similarity between a symbol and the thing it refers to has been shown to facilitate young children’s understanding of symbols [11]. The FPS has been shown to be adequately understood by children as young as four years of age [1], with satisfactory supportive data on validity, reliability, and good psychometric properties [2].

**The Coloured Analogue Scale** (CAS) [29] is a 145 mm long vertical wedge with a color-shape gradient from narrow light pink at the bottom (“no pain”) to wide deep red at the top (“most pain”). Pain ratings are made by moving a marker to the position which best corresponds to pain intensity level, whereby more intense pain is represented by a higher position on the wedge. The CAS has been found to be easy to administer and score, and to have satisfactory psychometric properties [29].

**The Pieces of Hurt Tool** [22,23] requires the child to choose from zero to four

tokens (e.g., poker chips) to show how many “pieces of hurt” correspond with his/her pain intensity. Each token represents a piece of hurt, and declining to select any tokens represents no pain. This concrete ordinal rating scale has adequate psychometric properties and is suitable for use with children as young as 4 years of age [8].

**The Finger Span Measure** [16] is a folk derivative of the analog scale where pain is estimated as a proportion of the maximum span between the thumb and index finger. The child was told that touching together the tips of the thumb and forefinger (on the same hand) represented ‘no pain’, while stretching the thumb and forefinger as far apart as possible represented ‘most pain’. To obtain finger span scores, a 20cm length of fabric taping was used to measure the linear distances (maximum span, and pain-estimate span) between the nail tips of the thumb and forefinger when the tape was pulled taut and straight across the gap. Markings were made on the tape to represent the child’s maximum span and the pain-estimate span, and for formal scoring purposes the tape was removed and relative distances recorded in millimeters. A ratio score was then calculated by dividing the pain-estimate span by the maximal span, with scores thus ranging between 0 and 1. Although the psychometric properties of this measure are not well established, the Finger Span measure was included in the current study given its simple and apparently child-friendly response format [17].

**The Adjectival Scale** is an ordinal verbal descriptor scale comprising orally presented categories of “no pain,” “a little bit,” “quite a lot” and “most pain possible”. These responses were scored 0, 1, 2 and 3 respectively. Although the psychometric

---

<sup>1</sup> The Faces Pain Scale - Revised <sup>20</sup> is now available, it is scored on a 0-to10 metric. At the time of data collection it was not yet published. It is now freely available for all research and clinical uses, with instructions in multiple languages ( [www.painsourcebook.ca](http://www.painsourcebook.ca) ) (Available online 3 December 2004).

properties of adjectival pain scales are not well known, these measures have been considered useful in contexts where the use of equipment is not possible, such as for over-the-phone assessments [30].

**The Sydney Animated Facial Expression scale (SAFE)** [19,24] was developed at the Sydney Children's Hospital, Randwick, Australia<sup>2</sup>, as an animated version of the Faces Pain Scale (FPS) [2]. On the screen of a portable computer, a single face can change expression in a smooth animated sequence from the "no pain" to the "most pain possible" anchor points of the FPS. The SAFE scale was developed by computing the series of images between each successive pair of faces on the original scale, thus converting the 7-point faces series into a series of 101 faces. The SAFE is administered by instructing the child to push the left or right arrow key on the computer keyboard to change the face to an expression which corresponds with his or her pain intensity. The face at which the child stops the animation series can be recorded (via a separate keystroke by the investigator) as a numerical value from 0 to 100.

### *Procedure*

Following approval from the hospital ethics committee, written consent was obtained from parents and verbal assent from children. Children were then asked to identify a pain associated with their current episode of hospitalization, and to rate the intensity of this nominated pain experience using each of the six pain scales (FPS, CAS, Pieces of Hurt Tool, Adjectival Scale, Finger Span, and the SAFE). The scales were administered in random order according to a table of six-digit numbers generated by dice throw prior to enrolment of children into the study. After completing each pain rating,

---

<sup>2</sup> For further enquiries regarding the SAFE, contact Tiina Piira at the Pain Research Unit, Sydney Children's Hospital, Randwick, NSW 2031, Australia. PiiraT@sesahs.nsw.gov.au

children were asked to indicate the point (value) on each scale at which they “think a nurse should be called to help”. After using all six scales in this way, the children were asked to rank-order the six pain scales from the easiest to use through to the most difficult to use.

## Results

Interscale correlations for pain intensity ratings are listed in Table 1. The pain intensity ratings obtained with each of the scales were found to be highly positively inter-correlated with ratings obtained using each of the other five scales ( $r = .53$  to  $r = .75$ , all  $p < .001$ ). Supporting the proposition that the scales were measuring the same construct, Cronbach’s alphas reflecting agreement across all six scales were 0.90 for younger children and 0.92 for older children.

*[Insert Table 1 about here.]*

Table 2 lists the raw means and standard deviations for each of the six pain self-rating scales for younger and older children. To facilitate comparison, pain scores were standardized using a linear transformation to a common metric of 0-10 [34,35]. A within-subject ANOVA was conducted to determine whether scores from the six instruments differed significantly from one another. This revealed a large overall effect of scale ( $F = 21.60$ ,  $p < .001$ ). Paired t-tests were conducted to compare the means for each scale to the grand mean of all six standardized pain intensity scores ( $\bar{X} = 6.13$ ,  $SD = 2.21$ ;  $\bar{X} = 5.29$ ;  $SD = 1.84$  for younger and older children respectively). Only the FPS and the Finger Span measures did not differ significantly from the grand mean for both younger (FPS:  $t$

= -.73,  $p = .47$ ; Finger Span:  $t = 1.27$ ,  $p = .21$ ) and older children (FPS:  $t = -.29$ ,  $p = .77$ ; Finger Span:  $t = -.53$ ,  $p = .60$ ). These mean differences between measures should be interpreted with caution as some differences may be an artifact of the smaller number of response options on the Pieces of Hurt and Adjectival scales. An additional reason for caution in interpreting the latter analysis is that it depends on an assumption that each scale is a linear interval scale, which may not be justified.

Table 3 lists the raw and standardized mean pain ratings for the point at which children thought that “a nurse should be called to help”. While the pain intensity ratings at which point older children felt that a nurse should be called for help across the six measures were fairly similar, there was more variability in these scores for younger children. In particular, younger children tended to give higher scores when using the Pieces of Hurt, and lower scores when using the CAS.

*[Insert Table 3 about here.]*

There were 65 children who were able to provide a meaningful preferential rank ordering of the scales. Missing data for these rankings was due to some children (mostly younger children) being unable to understand the task required. Rank-orderings of the scales across age are listed in Table 4. The SAFE was rated as the easiest to use, followed by the FPS and the CAS. The Adjectival Scale and Finger Span were rated as the hardest measures to use.

*[Insert Table 4 about here.]*

To compare the measures with respect to possible bias and skew in ratings, the response distributions for younger and older children on each of the six scales were examined. These response distributions are presented in Figure 1, which shows the percentage of children using each scale who provided a relatively low, medium or high band of score. The bands of low, intermediate and high reflect categories of pain intensity ratings on the 0-10 scale in purely quantitative terms, rather than qualitatively distinct pain experiences per se: “low” describes scores from 0.0 to 3.33, “medium” refers to scores from 3.34 to 6.67, and “high” refers to scores from 6.68 to 10. Tests of skewness from a normal distribution were not statistically significant for any measure in either age group.

[Insert Figure 1 about here.]

When multiple measures are available for the same construct, it has been proposed that the best estimate of that construct is a composite of the separate scores [32]. One strategy for identifying the best single-score estimate for the composite construct score is to calculate an item-total correlation. In the present context, this correlation, for each scale, will reflect the strength of association between the pain rating using that specific measure and a ‘total’ score reflecting the composite of ratings using the remaining five measures. This method was applied to the current set of pain ratings applying the Spearman-Brown correction, and the resulting rank order of these corrected item-total correlations is listed in Table 3. In descending order of magnitude, the corrected item-total correlation coefficients were: CAS .82; FPS .81; Adjective .77;

Pieces of Hurt .75; SAFE .74; Finger Span .70, indicating that there was moderately high convergent validity across the measures.

*[Insert Table 4 about here]*

A rank ordering of the six measures according to the degree of relative bias of the distribution of scores toward the upper end (i.e., 'most pain' anchor point) of each scale is provided in Table 4. Compared with the SAFE, FPS and CAS, response distributions for the Pieces of Hurt tool, Adjectival, and Finger Span scales showed the following pattern: a greater proportion of children selecting ratings from the top third of the scale, with a smaller proportion selecting ratings from the bottom third of the scale. This pattern of responses was particularly true for the younger children in this sample.

If the three parameters listed in Table 4 (preference, item-total correlation, and response distribution) are weighted equally, then the two optimum measures are seen to be the FPS and the CAS, as these two scales have the lowest rank sum.

### **Discussion**

This study compared six self-report measures of children's pain intensity to provide information on the cross-calibration of these instruments with respect to ratings of the same pain episode. All six measures of pain intensity were found to be quite highly correlated with each other.

When comparing standardized scores for younger children (between 4 and 8 years), some scales produced relatively higher average scores than others. Notably, the Pieces of Hurt tool resulted in a somewhat higher mean pain intensity score than the other

measures. Similar findings with the Pieces of Hurt tool have been reported in earlier studies with healthy children undergoing routine immunization via intramuscular injection [18].

In rating the level of pain at which point younger children thought a nurse should be called to help, the Pieces of Hurt tool again produced comparatively higher ratings than other measures of pain intensity. By way of contrast, for older children, the standardized scores of pain intensity across each of the six measures were quite similar, as were their ratings of the point at which they thought that a nurse should be called.

The greater variation in scores across scales for younger children, as compared with older children, may be due to younger children being more influenced by specific scale characteristics, whereas older children may be more readily able to map their pain experience onto a wider variety of response formats. Alternatively, it is possible that younger children may have made relatively more random responses than older children.

In considering the variance in pain ratings across different pain scales, the notion of a “true pain” score seems problematic and of limited utility. Rather it has been suggested that pain ratings exist within a social and contextual framework [38]: an individual communicates pain within the constraints of the tools provided for the assessment and the individual’s appraisal of how the ratings will be interpreted and used. It is imperative for the researcher and clinician alike to have some understanding of this communication process.

The issue of whether a certain pain score indicates a need for treatment has often been discussed in the context of clinical policy development. However, given the significant individual differences which exist in children’s pain ratings, and the

differences resulting from the choice of pain assessment tool used, such an approach does not seem entirely feasible. Remaining cognizant of these issues, data from the present study revealed that the children thought that, on average, their pain needed to be addressed, and a nurse called to help, for scores approximating 4 or 5 on a 0-to-10 pain intensity scale. Certainly this would suggest that further assessment and consideration of pain management should occur at this point; however, a decision to intervene should be made in the context of other relevant information. Rigorous protocols whereby interventions are initiated at certain pain intensity scores may take attention away from the many contextual factors which influence children's self-reports of pain. These include the fear of being given a needle or medicine, or the possible reinforcement of pain behaviors [8].

In addition to psychometric issues, selection of scales for use with children may also reflect factors such as the formats preferred by children [8]. In the present sample, both younger and older children preferred the animated facial expression scale (the SAFE), rating it as the easiest scale to use, followed by the FPS and CAS. Notably these three measures also showed evidence for relatively less upper end bias than the other scales. As with other applications of computer technology to the pediatric pain self-report [4], it is likely that most of the children in the sample were at ease with use of the computer and may have enjoyed the novelty of the SAFE relative to the other non-computer measures. Moreover, the fact that children reported that the two facial expression scales were easiest to use is consistent with other studies reporting facial expression measures of pain intensity to be appealing to children [15,36,39]. Unlike some

other pain assessment scales, facial scales do not rely heavily on language, and most children can immediately and intuitively relate to the instrument [25].

The Adjectival and Finger Span measures were rated by children as the hardest to use. It may be difficult for children, particularly young children, to relate their pain experience to verbal descriptors (as required in the Adjectival scale), especially if the words do not correspond with a personal pain vocabulary conveying incremental intensity. Although the Finger Span measure is less reliant on language, and is based on a type of gesture that is familiar to most children, the present sample apparently found it difficult to relate pain experience to this type of response format. Some older children in the sample did comment that this scale type was a ‘baby’ gesture.

The present study included pain scores collected at only one occasion for each child. Therefore the current data do not address the validity and reliability of each pain measure for detecting meaningful changes in pain over time, or in response to intervention (e.g., the Minimum Clinically Significant Difference) [3]. Future calibration studies could include repeated measures using the most promising scales. Such data will help to identify which scales are most sensitive to pain-relieving or pain-producing interventions for different age groups.

In using multiple measures to assess a single construct there is the possibility that the raters are influenced by their memory of how they rated the construct on the previous measures. However, this potential confounding is kept to a minimum through the use of scales with very different response formats, and by presenting the scales in various random orders.

This study was heterogeneous with respect to the type of pain rating. First, there was some degree of temporal variance, i.e., either current or recalled pain, depending on whether the child had pain at the particular moment of data collection. Although all participants had experienced periods of pain while in hospital, many were not in pain at the time the researcher visited, due either to effective analgesia or to healing. The project targeted differences in children's reactions to different pain scales, but this rests on the assumption that a child consistently recollected and rated the same pain for each of the six scales. Future research might ideally include only currently experienced pain, although such an approach may restrict participation to children experiencing relatively mild pain who are able to attend to the tasks required.

Second, there was heterogeneity of pain type. While all participants were hospital inpatients, the study did not assess the possible impact of other non-pain distress such as fatigue, nausea and anxiety. The study also focused on pain intensity and did not solicit ratings on the distress or intrusiveness associated with a specific pain in a specific context. This dimension may be an important component of variance in opinion as to the "clinically significant" value on a scale. To this end, future research on children's appraisal of when a nurse should be called to help with pain could examine whether such ratings differ according to pain type, and affective impact of pain, and concurrent symptom load.

The present study did not include the popular Numerical Rating Scale (NRS), which is readily used by children about 8 years and older [13], requires no equipment, and commonly uses the 0-to-10 metric. A future study could incorporate the best measures from the present study in comparison with the NRS for older children.

In summary, given that each of the six measures of pain intensity was highly correlated with the others, use of any of these scales in a pediatric clinical context will provide useful information. It is important, however, to be mindful of the likelihood that some measures may yield consistently higher or lower scores, particularly in young children. Children's scale preferences should also be taken into consideration when selecting a pain intensity measure, if a choice is possible. The present study found that both younger and older children ranked facial expression measures as easier to use than other scale options. The FPS and the CAS each gave the highest correlations with the composite score based on all scales (see item-total correlations in Table 3). Moreover the FPS and the Finger Span measures were the only measures to provide mean scores which were not significantly different from the grand mean for the six pain measures. Although the management of pediatric pain has not matured to the point that a single international standard has emerged for clinicians as the preferred self-report measure, on the basis of preference, internal consistency, response distribution, and low equipment cost, the FPS and CAS are recommended. It should be noted that since the time of testing for the current study, the FPS has been revised and the Faces Pain Scale – Revised (FPS-R) [24] is now more widely used around the world. The FPS-R has psychometric characteristics that are highly similar to the FPS [24] but holds the advantage of conforming to the widely popular 0-10 metric [35]. The CAS, other visual analog derivatives, and numerical rating scales are particularly suitable for older children. As the SAFE requires a dedicated portable computer and appears to have no important psychometric advantage over less expensive scales, it will probably be restricted to use in research rather than clinical contexts.

Future studies in this line of research, addressing some limitations of the current study, could be designed (a) to incorporate measures of clinical utility (e.g., time taken for training and administration; nurse and parent preference; cost and ease of dissemination); (b) to compare and cross-calibrate the Numerical Rating Scale with the CAS and FPS-R; and (c) to assess current rather than remembered pain, on repeated occasions rather than a single occasion, so as to permit assessment of retest reliability, pain trajectories, and response to intervention.

### **Acknowledgement**

This project was supported in part by a charitable grant from the James Kirby Foundation, and by the generous fund-raising efforts for the Pain Research Unit by Gay Windeyer and her Paediatric Pain Fund Raising Committee. Preliminary reports of the data were presented at the 5<sup>th</sup> International Symposium on Paediatric Pain, London, UK, November 2000. The participation of the third author was facilitated by two sabbatical travel grants from the University of Saskatchewan. We thank Robert Kovacs for assistance in developing the SAFE scale.

## References

1. Arts, S.E., AbuSaad, H.H., Champion, G.D., Crawford, M.R., Fisher, R.J., Juniper, K.H. & Ziegler, J.B. Age-related response to Lidocaine-Prilocaine (EMLA) emulsion and effect of music distraction on the pain of intravenous cannulation. *Pediatrics* 93: 797-801, 1994.
2. Bieri, D., Reeve, R.A., Champion, G.D., Addicoat, L., & Ziegler, J.B. The Faces Pain Scale for the self-assessment of the severity of pain experienced by children: development, initial validation, and preliminary investigation for ratio scale properties. *Pain* 41: 139-150, 1990.
3. Bulloch, B. & Tenenbein M. Assessment of clinically significant changes in acute pain in children. *Clinical Practice* 2002, 9: 199-202.
4. Calam R.M., Jimmieson P., Cox A.D., Glasgow DV, Larsen S.G. Can computer-based assessment help us understand children's pain? *European Journal of Anaesthesiology* 2000, 17: 284-288.
5. Chambers, C. & Craig, K. An intrusive impact of anchors in children's faces pain scales. *Pain* 1998; 78: 27-37.
6. Chambers, C.T., Giesbrecht, K., Craig, K.D., Bennett, S.M. & Huntsman, E. A comparison of faces scales for the measurement of pediatric pain: children's and parents' ratings. *Pain* 83: 25-35, 1999.
7. Chambers, C.T. & Johnston, C. Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology* 27: 27-36, 2002.

8. Champion, G.D., Goodenough, B., von Baeyer, C.L. & Thomas, W. Measurement of pain by self-report. In: Finley, G. and McGrath, P. eds., *Measurement of pain in infants and children*, Vol. 10. IASP Press, Seattle, WA, 1998, pp. 123-160.
9. Cohen, J. & Cohen, P. *Applied multiple regression/correlation analysis for the behavioral sciences* (2<sup>nd</sup> Edn., pp33-34). Hillsdale, NJ: Lawrence Erlbaum, 1983.
10. Craig, K.D., Lilley, C.M. & Gilbert, C.A. Social barriers to optimal pain management in infants and children. *Clin J Pain*. 12: 232-242, 1996.
11. DeLoache, J.S. Dual representation and children's use of scale models. *Child Development* 71: 329-338, 2000.
12. Demyttenaere, S., Finley, G.A., Johnston, C.C. & McGrath, P.J. Pain treatment thresholds in children after major surgery. *Clin J Pain*. 17: 173-7, 2001.
13. Fanurik, D., Koh, J.L., Harrison, R.D., Conrad, T.M. & Tomerlin, C. Pain assessment in children with cognitive impairment. An exploration of self-report skills. *Clin Nurs Res*. 7: 103-19, 1998.
14. Finley, G.A., McGrath, P.J., Forward, S.P., McNeill, G. & Fitzgerald, P. Parents' management of children's pain following 'minor' surgery. *Pain* 64: 83-87, 1996.
15. Fogel-Keck, J, Geckensmeyer, J.E., Joyce, B.A. & Schade, J.G. reliability and validity of the faces and word descriptor scales to measure procedural pain. *Journal of Pediatric Nursing* 11: 368-374, 1996.

16. Franzén, O.G. & Ahlquist, M.L. The intensive aspect of information processing in the intradental A-delta system in man: a psychophysiological analysis of sharp dental pain. *Behaviour and Brain Research* 33: 1-11, 1989.
17. Gaffney, A., McGrath, P.J. & Dick, B. Measuring pain in children: developmental and instrumental issues. In: N.L. Schechter, C.B. Berde and M. Yaster eds., *Pain in Infants, Children, and Adolescents* (2<sup>nd</sup> Ed). Lippincott Williams & Wilkins: Philadelphia. 2003.
18. Goodenough, B., Addicoat, L., Champion, G.D., McInerney, M., Young, B., Juniper, K., & Ziegler, J.B. Pain in 4- to 6-year-old children receiving intramuscular injections: a comparison of the Faces Pain Scale with other self-report and behavioral measures. *Clin J Pain* 13: 60-73, 1997.
19. Goodenough, B., Champion, G.D., von Baeyer, C.L. & Ziegler, J.B. Facial expression scales for self-report of pain by children: rationale, validity and utility. *Fourth International Symposium on Pediatric Pain*. Helsinki, Finland. July, 1997.
20. Goodenough B, Thomas, W., Champion, GD et al. (1999). Unravelling age effects and sex differences in needle pain: ratings of sensory intensity and unpleasantness of venipuncture pain by children and their parents. *Pain*, 80: 179-190
21. Hamers, J.P., Abu-Saad, H.H., van den Hout, M.A., Halfens, R.J. & Kester, A.D. The influence of children's vocal expressions, age, medical diagnosis and information obtained from parents on nurses' pain assessments and decisions regarding interventions. *Pain* 65: 53-61, 1996.

22. Hester, N. The preoperational child's reaction to immunization. *Nursing Research* 28: 250-255, 1979.
23. Hester, N.O., Foster, R. & Kristensen, K. Measurement of pain in children: generalizability and validity of the Pain Ladder and the Poker Chip Tool. In: Tyler, D.C. and Krane, E.J. eds. *Advances in Pain Research and Therapy: Pediatric Pain*, Vol. 15. Raven Press, New York, 1990, pp. 79-84.
24. Hicks, C.L., von Baeyer, C.L., Spafford, P., van Korlaar, I., & Goodenough, B. The Faces Pain Scale - Revised: Toward a common metric in pediatric pain measurement. *Pain* 93: 173-183, 2001. Scale available at [www.painsourcebook.ca](http://www.painsourcebook.ca) (available at 3 December 2004)
25. Kuttner, L & LePage, T. Faces scales for the assessment of pediatric pain: a critical review. *Canadian Journal of Behavioral Science* 21:198-209, 1989.
26. Ling, M. & Anderson, C.T.M. Managing pediatric pain. *Practical Pain Management* Jan/Feb: 9-14, 2002.
27. McCaffery, M.& Pasero, C. *Pain: Clinical Manual*. Mosby, Toronto, 1999.
28. McGrath, P.A. *Pain in Children: Nature, Assessment, and Treatment*. Guilford, New York, 1990.
29. McGrath, P.A., Seifert, C.E., Speechley, K.N., Booth, J.C., Stitt, L., & Gibson, M.C., A new analogue scale for assessing children's pain: an initial validation study, *Pain* 64: 435-443, 1996.

30. Maio, R.F., Garrison, H.G., Spaite, D.W., Desmond, J.S., Gregor, M.A., Stiell, I.G., Cayten, C.G., Chew, C.L., Mackenzie, E.J., Miller, R.R. & Malley, P.J. Emergency Medical Services Outcome Project (EMSOP): pain measurement in out-of-hospital outcomes research. *Annals of Emergency Medicine* 40: 172-9, 2002.
31. Morton, N.S. Simple and Systemic Management of Postoperative Pain. In G. Allen Finley and Patrick J. McGrath's (Eds.) *Acute and Procedure Pain in Infants and Children* 2001. IASP Press, Seattle.
32. Nunnally, C.J. *Psychometric Theory*. (2<sup>nd</sup> Edn). 1978. McGraw-Hill, New York.
33. Svensson, E. Concordance between rating using different scales for the same variable. *Statistics in Medicine* 19: 3483-3496, 2000.
34. Tesler, M.D., Savedra, M.C., Holzemer, W.L., Wilkie, D.J., Ward, J.A. & Paul, S.M. The Word-Graphic Rating Scale as a measure of children's and adolescents' pain intensity. *Research in Nursing and Health* 14: 361-371, 1991.
35. von Baeyer, C.L., & Hicks, C.L. Support for a common metric for pediatric pain intensity scales. *Pain Research and Management* 5: 157-160, 2000.
36. West, N., Oakes, L., Hinds, P.S., Sanders, L., Holden, R., Williams, S., Fairclough, D., & Bozeman, P. Measuring pain in pediatric oncology ICU patients. *Journal of Pediatric Oncology Nursing* 11: 64-68, 1999.
37. Westman HR. Postoperative complications and unanticipated hospital admissions. *Seminars in Pediatric Surgery* 8: 23-9, 1999.

38. Williams, A.C.deC, Davies, H & Chadury, Y., Simple pain rating scales hide complex idiosyncratic meanings. *Pain* 85: 457-463, 2000.
39. Wong, D.L. & Baker, C.M. Pain in children: Comparison of assessment scales. *Pediatric Nursing* 14: 9-17, 1988.
40. Woodgate, R. & Kristjanson, L.J. A young child's pain: how parents and nurses 'take care'. *International Journal of Nursing Studies* 33: 271-284, 1996.

**Table 1.** Intercorrelations between the child ratings for the six pain intensity measures.

	Faces Pain Scale	Coloured Analogue	Pieces of Hurt †	Adjectival †	Finger Span
<b>Coloured Analogue</b>	.75*				
Pieces of Hurt †	.62*	.69*			
Adjectival †	.65*	.73*	.70*		
Finger Span	.57*	.62*	.70*	.60*	
SAFE	.72*	.69*	.53*	.59*	.56*

\*  $p < .001$ 

† As the Adjectival and Pieces of Hurt measures are ordinal scales, Spearman-Brown correlations have been reported. All other correlations in the above table are Pearson correlations.

**Table 2.** Means (standard deviations) for each of the six pain intensity measures for younger (4-7 years) and older (8-16 years) children, showing pain intensity ratings for rated pain episode.

Scale	Raw Pain Intensity rating (SD)		Standardized Pain Intensity Ratings on 0-10 scale (SD)	
	Younger (4-7yrs) (n = 37)	Older (8-16yrs) (n = 45)	Younger (4-7yrs) (n = 37)	Older (8-16yrs) (n = 45)
Faces Pain Scale (0-6)	3.56 (1.70)	3.13 (1.46)	5.93 (2.83)	5.22 (2.43)
<b>Coloured Analogue (0-10)</b>	5.44 (3.07)	4.66 (2.21)	5.44 (3.07)	4.66 (2.21)
Pieces of Hurt (0-4)	2.83 (0.94)	2.19 (0.89)	7.08 (2.35)	5.47 (2.22)
Adjectival (0-3)	1.83 (0.74)	1.64 (0.65)	7.08 (1.84)	6.61 (1.61)
Finger Span (0-1)	0.66 (0.29)	0.52 (0.21)	6.59 (2.91)	5.17 (2.08)
SAFE (0-100)	50.37 (29.51)	45.96 (24.73)	5.04 (2.95)	4.60 (2.47)

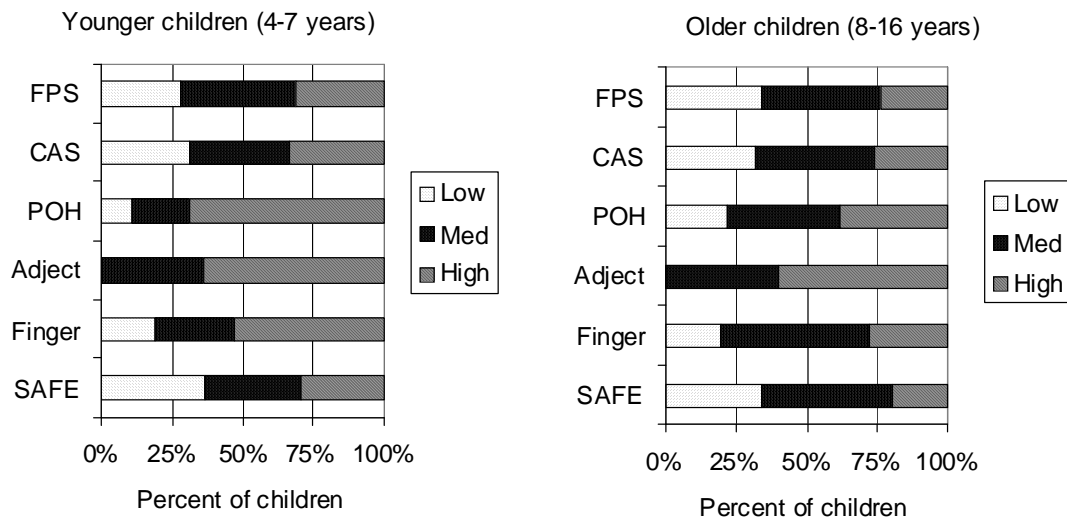
**Table 3.** Means (standard deviations) for each of the six pain intensity measures for younger (4-7 years) and older (8-16 years) children, showing the point at which children thought that “a nurse should be called”.

Scale	Raw Mean Score (SD)		Standardized Mean Score (SD)	
	Point at which to call nurse		Point at which to call nurse	
	Younger (4-7yrs) (n = 37)	Older (8-16yrs) (n = 45)	Younger (4-7yrs) (n = 37)	Older (8-16yrs) (n = 45)
Faces Pain Scale (0-6)	2.58 (2.04)	3.21 (1.09)	4.30 (3.39)	5.36 (1.82)
<b>Coloured Analogue (0-10)</b>	3.79 (2.28)	4.71 (1.72)	3.79 (2.28)	4.71 (1.72)
Pieces of Hurt (0-4)	2.47 (1.12)	2.13 (0.68)	6.05 (2.54)	5.33 (1.70)
Adjectival (0-3)	1.56 (0.86)	1.59 (0.50)	5.19 (2.85)	5.32 (1.66)
Finger Span (0-1)	0.57 (0.29)	0.48 (0.20)	5.71 (2.93)	4.81 (2.05)
SAFE (0-100)	45.30 (32.93)	42.84 (21.90)	4.53 (3.29)	4.28 (2.19)

**Table 4:** Rank-order of scales (1=best) according to overall child preference, strength of correlation with composite pain rating (Item-Total correlation), and relative lack of bias in response distribution,.

Scale	Child preference	Item-total correlation	Lack of upper-end bias in response distribution
Faces Pain Scale	2	2	1
Coloured Analogue	3	1	2
Pieces of Hurt	4	4	6
Adjectival	6	3	5
Finger Span	5	6	4
SAFE	1	5	3

**Figure 1.** Distributions of pain scores on each scale expressed as the proportion of children whose rating fell within the bands of relatively low, intermediate and high pain intensity ratings. For this purpose, scores on each scale were standardized to a common metric of 0-10 (the reader is referred to Table 2 for raw scores), and “Low” refers to scores from 0.0 to 3.33 out of 10, “Med” refers to intermediate scores from 3.34 to 6.67, and “High” refers to scores from 6.68 to 10.0



FPS = Faces Pain Scale; CAS = Coloured Analogue Scale; POH = Pieces of Hurt tool; Adject = Adjectival measure; Finger = Finger Span measure; SAFE= Sydney Animated Facial Expression scale.