Validity and reliability assessment of the Brazilian version of the Faces Pain Scale–Revised

Claudia L. E. Charry, Joseane dos Santos Piola, Maria B. M. Linhares, and José Aparecido da Silva

Universidade de São Paulo, Ribeirão Preto, SP, Brazil

Abstract

The Faces Pain Scale–Revised (FPS–R) is among the most commonly used measures of pain intensity in clinical and research settings. Little evidence exists about the Brazilian version of this scale. The purpose of the present study was to examine the validity and reliability of the Brazilian version of the FPS–R. The sample comprised 214 children, 6 to 10 years of age. In the first phase of the study, the children ranked the faces of the FPS–R according to pain intensity, and the faces were presented in all possible pair combinations (content validity). All six faces were correctly ranked by 47% of the younger children (6 to 7 years old) and 52% of the older children (8 to 10 years old). With regard to paired combinations, all of the pairs were correctly placed by 63% of the younger children and 67% of the older children. In the second phase, the children rated recalled experiences of pain using both the Brazilian version of the FPS–R and Coloured Analogue Scale (CAS; convergent validity). The children were retested 4 weeks later (test–retest reliability). The degree of agreement between the FPS–R and CAS was moderate, with a Kendall's tau-b (τ) of .49 ($p < .01$). The test-retest reliability coefficient was $\tau = .52$ ($p < .01$). This study provides evidence of the validity and reliability of the FPS-R as a pain intensity measurement for use in Brazilian samples. **Keywords:** pain assessment, faces pain scale–revised, validity, reliability.

Introduction

Self-report methods are an important source of information for pediatric pain assessment (Bulloch & Tenenbein, 2002; McGrath, & Gillespie, 2001; O'Rourke, 2004). Several instruments are available in which children describe their subjective experiences of pain (Stinson, Kavanagh, Yamada, Gill, & Stevens, 2006). Among them, faces pain scales are one of the most frequent used. This type of scale is frequently applied to measure pain intensity. They are easy to understand by the children and quick to apply (Chambers & Craig, 1998; Chambers, Giesbrecht, Craig, Bennett, & Huntsman, 1999). One of the most recommended scales is the Faces Pain Scale–Revised (FPS–R; Hicks, von Baeyer, Spafford, van Korlaar, & Goodenough, 2001; Spagrud,

Piira, & von Baeyer, 2003), which has six different facial expressions that help measure pain intensity. This scale has important advantages such as the absence of smiles and tears that may confound affective states with pain ratings (Chambers & Craig, 1998; Chambers et al., 1999). It also offers the possibility of obtaining scores that are compatible with other scales that use a common metric such as 0-5 or 0-10 (Hicks et al., 2001).

Recent reviews of the literature on tools that evaluate pain intensity in children and adolescents included the FPS–R on the list of suggested instruments because of its psychometric properties and feasibility (Srouji, Ratnapalan, & Schneeweiss, 2010; Tomlinson, von Baeyer, Stinson, & Sung, 2010). In another evidence-based review of pediatric pain measures, Cohen et al. (2008) classified the FPS–R as a "well-established assessment," which has sufficient evidence and available information.

Several versions of the FPS–R are available in different languages (von Baeyer, Wood, & Piira, 2005). Some of these versions have been tested in countries like France (Wood, von Baeyer, Bourrillon, Dejos-Conant, Clyti, & Abitbol, 2004), Thailand (Newman, Lolekha, Limkittikul, Luangxay, Chotpitayasunondh, & Chanthavanich, 2005), and Spain (in both Spanish and Catalan; Miró, & Huguet, 2004; Miró, Huguet, Nieto, Paredes, & Baos, 2005). Instructions have been translated into Brazilian Portuguese and back-translated

Claudia L. E. Charry, Joseane dos Santos Piola, and José A. Da Silva, Departamento de Psicologia, Ribeirão Preto, Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, Brazil; Maria B. M. Linhares, Departamento de Neurociências e Ciências do Comportamento, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Brazil. Correspondence regarding this article should be directed to: Claudia Ligia Esperanza Charry Poveda, Department of Psychology, FFCLRP/USP, Av. Bandeirantes, 3900, Ribeirão Preto, São Paulo, 14040-901, Brazil. E-mail: claudialich@pg.ffclrp.usp.br

by Charry, Da Silva, Passareli, Santos, & Linhares (2007), and these versions are currently available (www. iasp-pain.org/fpsr; accessed September 2, 2013). An independent translation and adaptation was published by Silva & Thuler (2008).

The application of an instrument in another population that is different from the original population, especially one with a different language, must be preceded by an adaptation process. This process involves not only translation but also the evaluation of psychometric properties with regard to the population of interest (Miró & Huguet, 2004; Reichenheim & Moraes, 2007; Sousa & Rojjanasrirat, 2011; Kottner et al., 2011). The validity and reliability of an instrument cannot be assumed to be equal for different populations, especially when differences in language and culture are significant.

Little evidence exists of the psychometric properties of the FPS–R for use in Brazilian samples; for this reason the aim of the present study was to assess both the test-retest reliability and content and convergent validity of the Brazilian version of the FPS–R.

Methods

Participants

The study sample included 214 children, 113 girls and 101 boys (mean age, 8.31 years; standard deviation, 1.3) who were recruited from eight schools in Ribeirão Preto, Brazil (four public schools and four private schools). Inclusion criteria included the following: (*i*) age between 6 and 10 years, (*ii*) absence of apparent cognitive impairment, (*iii*) ability to understand and speak Portuguese, and (*iv*) ability to understand the use of the instruments of this study (i.e., the FPS–R and Coloured Analogue Scale [CAS]).

A total of 420 invitation letters was distributed to several of the 1st through 5th grade classrooms. The letter included a description of the study and a free informed consent form. Parents who were interested in allowing their child to participate in the study had to complete and return the form. Two hundred eighteen parents provided consent for their child to participate. Four children were excluded from the study because their age was incompatible with the inclusion criteria.

Apparatus

Faces Pain Scale–Revised (Hicks et al., 2001). The FPS–R is a self-rated and self-administered scale designed to measure a child's level of perceived pain intensity. It consists of six faces, presented horizontally, that depict different degrees of pain, from "no pain" to "most pain possible." A numerical value from 0 to 10 (i.e., 2, 4, 6, 8, 10) is assigned to each face. The child points to the face that shows the severity of the pain they experienced. The FPS–R has been tested in several samples under both clinical and nonclinical conditions (Migdal, Chudzynska-Pomianowska, Vause, Henry, & Lazar, 2005; Miró & Huguet, 2004; Newman et al., 2005; Wood et al., 2004). It has demonstrated reliability

and validity for measuring the severity dimension of pain.

Coloured Analogue Scale (McGrath, Seifert, Speechley, Booth, Stitt, & Gibson, 1996). The CAS is a visual analogue scale that measures the intensity of pain experienced. The CAS consists of a mechanical device with a horizontal marker that slides over a tetragon that varies in width and hue from narrow and white at the bottom (labeled "no pain") to wide and dark red at the top (labeled "most pain"). Each scale position has a numerical value, which is on the back of the scale (scored from 0 to 10 in increments of .25). The CAS has been shown to be a reliable and valid measure in different samples of children and adolescents (Bulloch & Tenenbein, 2002; McGrath et al., 1996; Piira, Hayes, Goodenough, & von Baeyer, 2006).

Procedure

The present study was approved by the Research Ethics Committee of the School of Philosophy, Sciences and Letters of Ribeirão Preto, University of São Paulo (no. 024/2007). Prior to data collection, parents signed the informed consent. The children received age-appropriate explanations about the study, and children who expressed verbal assent were included in the sample. The children were tested individually in a quiet room. The interviews lasted ~5 min each and took place during school hours.

The adaptation of the FPS–R instructions to Brazilian Portuguese (Charry et al., 2007) began with an evaluation of the text by four different professionals (a teacher, a nurse, a psychologist and a pediatrician) who were familiar with the language that children use. The evaluation was conducted to perform the necessary linguistic adjustments. Some minor modifications were made, and the final text was back-translated from Brazilian Portuguese to English by a bilingual professional who was naive about the original version of the instrument. The back-translated version was sent to the author of the original instrument who verified that it was appropriate. Finally, a pilot test was conducted with 31 school-aged children to check their understanding of the instructions and use of the FPS–R. No changes were made as a result of this pilot test because the instructions and use of the scale were clear for the children.

First phase. In this phase, the children performed one of two tasks to evaluate the content validity of the FPS–R. The two tasks were the following: Task 1 (to organize the faces for the expression of pain) and Task 2 (to compare all possible pairs of faces). The FPS–R was assumed to express a sequence that was understandable to children; thus, they could place the faces on the scale in the expected order; in the paired comparisons, they could show the face that expressed more pain.

In Task 1, each face of the FPS–R was mounted on 9.5 × 7.8 cm cardboard. Upon presentation of the faces, the children were given the following instruction: "I am going to show you different drawings of faces that express several pain degrees. I want you to help me

organize them. Start with the face you think expresses the worse pain and then the face that expresses a little less of pain, and so on until the last face that doesn't express any pain at all." The order was randomized. Some children were asked to organize the faces from the most to the least painful, and other children were asked to organize the faces from the least to the most painful. In Task 2, the faces of the FPS–R were presented in pairs. All possible combinations were shown randomly with the following instruction: "I am going to show you two drawings of faces. I want you to choose the face that expresses more pain."

Separate groups of children were established for each task. These tasks, the organization of faces, and comparisons by pairs were used by Bieri, Reeve, Champion, Addicoat, & Ziegler (1990) to validate the FPS. Altogether, 83 children participated in this first phase. Forty children participated in the first task, and 43 participated in the second task.

Second phase. The convergent validity and test-retest reliability of the Brazilian version of the FPS–R were evaluated. The interviewer asked the children to remember a recent painful experience and rate it using the FPS–R and CAS. The order of presentation was randomized. The CAS is an instrument that is also used to assess the intensity of pain and, therefore, is expected to have a significant degree of agreement between the two sets of measurements.

One hundred children participated in this phase. The data from nine children were excluded because they evaluated emotional pain experiences, and only physical pain experiences were considered for this study.

Four weeks later, the test-retest reliability of the Brazilian version of the FPS–R was tested. The children were instructed to recall the same personal pain experience and rate it again. These two different estimations at two points in time were expected to have an important degree of agreement. Retest data from five children were not available because they did not attend school during our second visit.

Statistical analysis

To examine age differences in the children's use of the scales, the participants were divided into two age groups: 6–7 years and 8–10 years. Associations between pain ratings and age were examined using Spearman's correlation coefficient. Percentages were used to explore the responses of children in the two tasks in the first phase. In the second phase, Kendall's tau-b statistic test was used to examine the degree of agreement between the FPS–R and CAS ratings and assess the test-retest reliability of the FPS–R. Possible values of Kendall's tau-b range from

-1 to +1, in which +1 indicates all pairs are ordered in the same way and -1 indicates they are all ordered in the opposite way. Additionally, an index of the proportion of discordant scores was calculated (i.e., scores that differed by more than 2/10 between the CAS and FPS–R). This is the minimum significant difference

that can be detected in the FPS–R (Tovar, von Baeyer, Wood, Alibeu, Houfani, & Arvieux, 2010).

Results

Content validity

Table 1 presents the percentages of children who ranked the faces in the expected order in both tasks: sequencing the six faces and using paired comparisons. The percentages of faces 0, 2, and 10 were higher than the middle faces on the ordinal scale (faces 4, 6, and 8) for both ages. The middle faces also presented a less homogeneous percentage than the extreme faces. The children had more difficulty ranking faces 4 and 6 in both tasks, ordering the six faces simultaneously, and making comparisons of all possible pairs. Face 6 showed the lowest percentage. When faces 4 and 6 were not included, the percentages increased significantly.

Table 1. Children's responses in Tasks 1 and 2 (percentages)

Age groups	Children who ranked the faces individually in the expected order (%) Faces						All faces in the expected order (%)	All faces in the expected order, excluding faces 4 and 6 (%)	All pairs in the expected order (%)
	0a	2	4	6	8	10			
6-7 years	100	94	59	47	76	94	47	71	63
8-10 years	96	91	61	52	83	100	52	78	67

aFace 0 represents no pain. Face 10 represents the most pain possible.

Convergent validity

The median pain intensity scores were 8 on the FPS–R (interquartile range, 6–10) and 7.75 on the CAS (interquartile range, 6–8.75; Figure 1). The degree of agreement between the FPS–R and CAS ratings was examined. Table 2 presents Kendall's tau-b statistic for FPS–R and CAS ratings. The values obtained were significant and positive.

Pain ratings were not significantly correlated with age on either the FPS–R (rs = .10, p = .32) or CAS (rs = .13, p = .20). The proportion of discordant scores between the FPS–R and CAS was also calculated (i.e., ratings that differed by more than 2/10 between the two scales). This proportion was 12% for the younger children (6 to 7 years old) and 15% for the older children (8 to 10 years old). A total of 14% of this sample gave discordant scores.

Table 2. Concordance measures (Kendall's tau-b) between the Brazilian version of the FPS–R and CAS ratings

Age groups	n	τ
6-7 years	33	.66
8-10 years	58	.37
Total	91	.49

All coefficients are statistically significant at p < .01.

Test-retest reliability

The reliability of the FPS–R was tested by asking children to recall a painful experience at two different moments in time, with a 4-week interval. The test-retest reliability coefficient was $\tau = .52$ ($p < .01$). This value suggests an acceptable temporal stability for this scale.

Discussion

The primary purpose of the present study was to establish the validity and reliability of the Brazilian version of the FPS–R. Our findings indicate that this scale has good psychometric properties and can be a suitable instrument for measuring pain intensity in Brazilian children.

Data for the first phase indicated that the Brazilian version of the FPS–R has appropriate content validity. The variation in the middle faces was higher than in the extreme faces, suggesting that faces 4 and 6 were more difficult for children to rank. The similarity of the features of the middle faces likely made discriminating the one that expressed more pain difficult. The percentages for younger children were slightly lower than for older children. This fact reinforces the necessity of taking special care when this type of scale is applied in younger children, considering both the number of response categories of the instrument and similarity among them (Decruynaere, Thonnard, & Plaghki, 2009; Hunter, McDowell, Hennessy, & Cassey, 2000).

The evaluation of reliability, specifically test-retest reliability, has some implicit difficulties, especially in the case of pain assessment instruments. Few studies have evaluated this psychometric property of the FPS–R. The findings in the present study indicated that the scores obtained using the FPS–R were relatively consistent over time. Moderate test-retest coefficients were obtained by Miró & Huguet (2004) who rated a list of hypothetical painful events with the Catalan version of the FPS–R.

Several studies have compared the FPS–R with other well-established pain intensity measures using either Pearson's or Spearman's correlation coefficient (Hicks et al., 2001; Miró et al., 2005; Silva, Thuler, & Leon-Casasola, 2011). Nevertheless, for some authors (Beckstead, 2011; Bland & Altman, 2010), comparisons between instruments that assess the same construct must be considered a case of the degree of agreement rather than a correlation. In the present study, the FPS–R showed moderate positive agreement with another scale, the CAS that assesses the same construct. Moderate-to-good Spearman's correlations were obtained by Newman et al. (2005) between the FPS–R and Visual Analogue Scale. Importantly, Kendall's tau-b values will always be less than Spearman's rank-order correlation coefficients (Sheskin, 2000). The total proportion of discordant scores obtained in this study falls within the limits of agreement suggested by von Baeyer (2012), in which at least 80% of the scores fall within 20% of the scale range. Large differences were found in concordance values between the two age groups, with poorer values in older children. This was an unexpected result and could correspond to the proportion of discordant scores, which was slightly higher in the older children. No similar data have been reported in other studies. This could be a result of difficulties using these scales, which is characteristic of this sample. More data are necessary to confirm this hypothesis.

The present study has some limitations. First, the participants evaluated an experience of recalled pain. This kind of task has been useful for the validation of pediatric pain scales (Bieri et al., 1990), but aspects such as memory and attention can affect children's self-reported pain intensity. Second, the children were in no pain or any clinical situation at the time of the evaluation. This characteristic has been presented in diverse studies of pain scales and represents a good alternative when the tasks require time or a specific disposition of the children (Bieri et al., 1990; Chambers & Craig, 1998; Hicks et al., 2001; McGrath et al., 1996). Unknown, however, is the degree to which our results may be generalized to clinical situations. Third, at the time when this research was conducted, no studies were published about translating the CAS instructions into Portuguese or validating this instrument for the Brazilian population. Few psychometric studies of this type of instrument have been performed in Brazil. The CAS is a recommended instrument that has excellent psychometric properties. It is well known in Brazil, appropriate for use with children between the ages of 6 and 10 years, and tested in different samples of children and adolescents. Nonetheless, specific data about its properties in the Brazilian population are unknown. Finally, in the present study, no-pain events were not evaluated by children, and the full ranges of the scales were not used. The agreement values were likely reduced by this restriction of the ranges of the scales.

Additional studies are needed to evaluate the psychometric properties of the Brazilian version of the FPS–R in children who experience clinical pain (i.e., postoperative pain and procedural pain) and its sensitivity (i.e., the scale's ability to detect clinically meaningful changes). Increasing the evidence of clinically significant differences between this instrument and other instruments and the interpretability of its scores will be important.

Acknowledgements

References

Beckstead, J. W. (2011). Agreement, reliability and bias in measurement: commentary on Bland and Altman (1986, 2010). *International Journal of Nursing Studies*, *48*, 134-135.

Bieri, D., Reeve, R. A., Champion, G. D., Addicoat, L., & Ziegler, J. B. (1990). 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.

Bland, J. M., & Altman, D. G. (2010). Statistical methods for assessing agreement between two methods of clinical measurement. *International Journal of Nursing Studies*, *47*, 931-936.

Bulloch, B., & Tenenbein, M. (2002). Validation of 2 pain scales for use in the pediatric emergency department. *Pediatrics*, *110*, e33.

Chambers, C. T., & Craig, K. D. (1998). An intrusive impact of anchors in children's faces pain scales. *Pain*, *78*, 27-37.

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 parent's ratings. *Pain*, *83*, 25-35.

Charry, C., Da Silva, J. A., Passareli, P., Santos, J., & Linhares, M. B. M. (2007, June). Instructions of FPS-R in Portuguese (Brazil). Retrieved from www.iasp-pain.org/fpsr; accessed September 2, 2013.

Cohen, L. L., Lemanek, K., Blount, R. L., Dahlquist, L. M., Lim C. S., Palermo, T. M., McKenna, K. D., & Weiss, K. E. (2008). Evidence-based assessment of pediatric pain. *Journal of Pediatric Psychology*, *33*, 939-955.

Decruynaere, C., Thonnard, J. L., & Plaghki, L. (2009). How many response levels do children distinguish on faces scales for pain assessment? *European Journal of Pain*, *13*, 641-648.

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

Hunter, M., McDowell, L., Hennessy, R., & Cassey, J. (2000). An evaluation of the Faces Pain Scale with young children. *Journal of Pain and Symptom Management*, *20*, 122-129.

Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B. J., Hrobjartsson, A., Roberts, C., Shoukri, M., & Streiner D. L. (2011). Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *International Journal of Nursing Studies*, *48*, 661-671.

McGrath, P., & Gillespie, J. (2001). Pain assessment in children and adolescents. In D. C. Turk, & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 97-118). New York: Guilford Press.

McGrath, P. A., Seifert, C. E., Speechley, K. N., Booth, J. C., Stitt, L., & Gibson, M. C. (1996). A new analogue scale for assessing children's pain: An initial validation study. *Pain*, *64*, 435-443.

Migdal, M., Chudzynska-Pomianowska, E., Vause, E., Henry, E., & Lazar, J. (2005). Rapid, needle-free delivery of lidocaine for reducing the pain of venipuncture among pediatric subjects. *Pediatrics*, *115*(4), e393-e398.

Miró, J., & Huguet, A. (2004). Evaluation of reliability, validity, and preference for a pediatric pain intensity scale: The Catalan version of the Faces Pain Scale-Revised. *Pain*, *111*, 59-64.

Miró, J., Huguet, A., Nieto, R., Paredes, S., & Baos, J. (2005) Valoración de la escala de dolor de caras–revisada (Faces Pain Scale–Revised) para evaluar la intensidad del dolor pediátrico en niños castellanos parlantes. *Revista de la Sociedad Española del Dolor*, *12*(7), 407-416.

Newman, C. J., Lolekha, R., Limkittikul, K., Luangxay, K., Chotpitayasunondh, T., & Chanthavanich, P. (2005). A comparison of pain scales in Thai children. *Archives of Disease in Childhood*, *90*, 269-270.

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

Piira, T., Hayes, B., Goodenough, B., & von Baeyer, C. L. (2006). Effects of attentional direction, age, and coping style on cold-pressor pain in children. *Behaviour Research and Therapy*, *44*, 835-848.

Reichenheim, M. E., & Moraes, C. L. (2007). Operationalizing the cross-cultural adaptation of epidemiological measurement instruments. *Revista de Saúde Publica*, *41*, 665-673.

Sheskin, D. (2000). *Handbook of parametric and nonparametric statistical procedures*, 2nd edition. Boca Raton: Chapman & Hall/ CRC.

Silva, F. C., & Thuler, L. C. (2008). Cross-cultural adaption and translation of two pain assessment tools in children and adolescents. *Jornal de Pediatria*, *84*, 344-349.

Silva, F. C., Thuler, L. C. S., & de Leon-Casasola, O. (2011). Validity and reliability of two pain assessment tools in Brazilian children and adolescents. *Journal of Clinical Nursing*, *20*, 1842-1848.

Sousa, V., & Rojjanasrirat W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, *17*, 268-274.

Spagrud, L., Piira, T., & von Baeyer, C. L. (2003). Children's self-report of pain intensity: The Faces Pain Scale–Revised. *American Journal of Nursing*, *103*, 62-64.

Srouji, R., Ratnapalan, S., & Schneeweiss, S. (2010). Pain in children: Assessment and nonpharmacological management. *International Journal of Pediatrics*, *2010*, 474838.

Stinson, J., 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 trails in children and adolescents. *Pain*, *125*, 143-157.

Tovar, C., von Baeyer, C. L., Wood, C., Alibeu, J.P., Houfani, M., & Arvieux, C. (2010). Postoperative self-report of pain in children: Interscale agreement, response to analgesic, and preference for a faces scale and a visual analogue scale. *Pain Research Management*, *15*, 163-168.

Tomlinson, D., von Baeyer, C. L., Stinson, J. N., & Sung, L. (2010). A systematic review of faces scales for the self-report of pain intensity in children. *Pediatrics*, *126*, e1168-e1198.

von Baeyer, C. L. (2012). Reported lack of agreement between self-report pain scores in children may be due to a too strict criterion for agreement. *Pain*, *153*, 2152-2153.

von Baeyer, C. L., Wood, C., & Piira, T. (2005). Instructions for administering the Faces Pain Scale–Revised in languages other than English. Retrieved from http://www.iasp-pain.org/fpsr; accessed September 2, 2013.

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.