Pain Medicine

Pain Medicine 2015; 16: 1982–1992
Wiley Periodicals, Inc.

METHODOLOGY, MECHANISMS & TRANSLATIONAL RESEARCH SECTION

Original Research Article

AN APP for the Assessment of Pain Intensity: Validity Properties and Agreement of Pain Reports When Used with Young People

Elisabet Sánchez-Rodríguez, MSc, Rocío de la Vega, PhD, Elena Castarlenas, PhD, Roman Roset, MSc, and Jordi Miró, PhD

Department of Psychology and Institut d'investigació Sanitària Pere Virgili, Unit for the Study and Treatment of Pain - ALGOS, Research Center for Behavior Assessment (CRAMC), Universitat Rovira I Virgili, Tarragona, Spain

Reprint requests to: Jordi Miró, PhD, Departament de Psicologia, Universitat Rovira i Virgili, 43007 Tarragona, Spain. Tel: (+34)977558179; Fax: (+34)977558880; E-mail: jordi.miro@urv.cat.

Conflict of interest: The authors have no conflicts of interest.

Abstract

Objective. *Painometer* is a mobile application that includes four pain intensity scales: the Numerical Rating Scale, the Faces Pain Scale-Revised, the mechanical visual analogue scale and the Colored Analogue Scale. The aim of this study was to analyze the validity and agreement of the intensity reports provided by these scales and their traditional counterparts.

Methods. Participants were 180 young people (mean age = 14.88; SD= 1.64; age range: 12–19). They were asked to report the maximum intensity of their most frequent pain in the previous three months using traditional and electronic versions of the scales. They also reported their level of fatigue and pain catastrophizing. Construct validity was evaluated by confirmatory factor analysis (CFA) and by convergent and discriminant validity. Criterion validity was assessed as concurrent validity. Agreement was calculated using the Bland and Altman method. Analyses were conducted for two confidence intervals (CI): 95% and 80%.

Results. CFA demonstrated that the four electronic versions of the scales measure a single factor. All the scales showed a) moderate to high convergent validity, b) adequate discriminant validity with fatigue ratings, and c) adequate concurrent validity with pain catastrophizing ratings. Results also show that traditional and electronic versions of the four scales are in agreement, at least at the 80% CI.

Conclusions. Our results demonstrate that pain intensity scores reported with the scales in *Painometer* are valid, and concordant with their traditional counterparts.

Key Words. Pain Intensity; Assessment; Mobile App; Smartphone; Young People

Introduction

Over the last 15 years, an increasing number of studies have used the so-called information and communication technologies (ICTs) for the management of pain in young people [1]. For example, internet-based interventions [2,3] are now available. Similarly, mobile devices such as PDAs [4–7] and Smartphones [8,9] are also used to assess young people with pain.

The use of ICTs to assess and treat people with pain has rapidly increased for a number of reasons, namely: a) they are more accessible and ubiquitous [10], b) data are captured in real-time [11], c) compliance is greater [12,13], d) accuracy is improved [4,14], and e) young people tend to prefer electronic devices to traditional test [7,8].

Although the fast development and increased use of ICTs for pain assessment is positive, there are some

caveats. One of the major disadvantages of digitalized questionnaires or self-report measures is that they are often used without previous rigorous analysis of their psychometrics properties and usability characteristics, as if the scores provided by them were just as suitable and valid as measures in a traditional format. Thus, although there is ample evidence that young people are capable of reliably reporting information about their pain [15,16], this is far from clear when digitalized versions of existing questionnaires are used, as only a handful of studies have looked into these issues, for example, [6–9,17,18]. Before these electronic versions of self-report measures can be extensively used, they must undergo close analysis.

Painometer is a mobile application that has been developed to help measure pain [17]. It contains the electronic versions of four widely used pain intensity scales: the Faces Pain Scale-Revised (FPS-R), the Numerical Rating Scale-11 (NRS-11), the Colored Analogue Scale (CAS), and the mechanical visual analogue scale (mVAS). Castarlenas et al. [8] studied the agreement between verbal and electronic versions of the NRS-11 included in *Painometer* in a sample of schoolchildren. They reported that pain intensity ratings on both scales seem to be comparable. However, it is yet to be determined whether the traditional administered versions and the electronic versions in *Painometer* behave similarly. The first major aim of this study was [1] to further analyze the psychometric properties of *Painometer*: that is to say, evaluate whether the four scales provided valid reports when used to measure young people's levels of pain intensity. Our hypothesis was that the electronic versions of the four pain intensity scales measure a single dimension and that they all provide valid reports when used to measure pain intensity in young people. The second major aim was to study the agreement between the reports provided by the four pain intensity scales contained in *Painometer* and their traditional counterparts. On the basis of published studies [7,8], we hypothesized that we would find an agreement between reports of the electronic and the traditional versions of the scales. A further secondary aim was to compare the participants' preference for traditional and electronic versions of the scales. On the basis of previous studies, we expected that participants would prefer the electronic versions.

Methods

Participants

Sample size analysis showed that 85 participants would be needed ($r = 0.3$; $\alpha = 0.05$; $\beta = 0.20$) to perform the planned analyses. On the basis of previous studies conducted by our research group, we predicted that about 40% of the participants would fail to return the signed consent. Thus, to have enough participants, we invited schoolchildren enrolled in grades 7 to 12 from two schools ($N = 280$).

A total of 203 schoolchildren (72.5%) returned the signed parental consent. Of these, 23 did not report having pain in the last 3 months and were excluded from the study. Additional exclusion criteria were 1) having a cognitive impairment and 2) not being able to understand Catalan or Spanish. Thus, the final sample was comprised of 180 schoolchildren, with 76 boys (42%) and 104 girls (58%). The participants were between 12- and 19-year old (mean = 14.88; SD = 1.636). All of them were Catalan native speakers, and all interviews and questionnaires were in Catalan. Table 1 summarizes descriptive information about the participants.

Measures

Pain intensity was measured with the 11-point Numerical Rating Scale, the Faces Pain Scale-Revised, the visual analogue scale and the Colored Analogue Scale. In all cases, participants had to report the maximum intensity of their most frequent pain in the last 3 months. In Painometer, instructions about how to use the electronic versions of each scale appear on the screen just before the scale so that users can read for themselves what to do and how to do it. However, in this study, instructions were verbally provided by the interviewer.

The 11-Point Numerical Rating Scale (NRS-11)

Participants reported the maximum intensity of their most frequent pain in the last three months on a scale from 0 (no pain) to 10 (very much pain). In its verbally administered version (vNRS-11) it has demonstrated good psychometric properties when used with children 6-year old or older [19–24]. The electronic version of the NRS-11 (eNRS-11) [17] was administered on an iPod Touch 4©. The iPod screen was 3.5 inches diagonally (or 88.9 mm). All the potential responses (i.e., numbers from 0 to 10) are depicted permanently at the top of the screen and the selected number is depicted in a bigger font in the middle of the screen. Participants could select the number that best represented their pain intensity in one of two different ways: 1) sliding their finger across the screen so that the eleven possible options could be seen by the user, one by one, in the center of the screen or 2) tapping on the chosen number at the top of the screen. Figure 1 shows a screenshot of the eNRS-11.

The Faces Pain Scale—Revised

FPS-R [25] is a self-report measure with six faces showing increasing pain intensity levels from left to right. As in the NRS-11, participants were asked to identify their most frequent pain in the last 3 months and select the face which best reflected the maximum intensity of this pain. The lower and upper anchors of the scale were "no pain" and "very much pain." FPS-R has been widely studied and shown good psychometric properties [25–27].

Table 1 Participants' descriptive information

Participants (N)	180
Mean age (SD)	14.88 (1.64)
Gender N (%)	76 (42%)
Boys	104 (58%)
Girls	
Pain status N (%)	112 (65%)
Acute pain	63 (35%)
Chronic pain*	72 (40%)
With pain at the time of the interview	
Schooling grade N (%)	34 (19%)
Seventh	22 (12%)
Eighth	29 (16%)
Ninth	61 (34%)
Tenth	16 (9%)
Eleventh	18 (10%)
Twelfth	
Localization of the most frequent pain N (%)†	43 (24%)
	5 (3%)
Head (excluded face)	8 (4%)
Throat/neck	2 (1%)
Shoulder	1 (1%)
>Chest	1 (1%)
Elbow	1 (1%)
Forearm	4 (2%)
Wrist	22 (12%)
Hand	4 (2%)
Abdomen	2 (1%)
Hip	7 (4%)
Groin/pubic area	24 (13%)
Thigh	7 (4%)
Knee	7 (4%)
Calf	8 (4%)
Ankle	2 (1%)
Foot	18 (10%)
Upper back	4 (2%)
Mid back	
Low back	

* Defined as a pain in any area of their body (1) lasting for at least 3 months and (2) present at least once a month.

† Information missing in 10 cases (5%).

The electronic version of the FPS-R (eFPS-R) shows the six faces at the top of the screen and the selected face in the center of the screen. Participants selected the face that best represented the intensity of their pain by sliding their finger across the screen or by tapping on the chosen face at the top of the screen. Figure 2 shows a screenshot of the eFPS-R as it was presented.

The Visual Analogue Scale

VAS is a 100 mm horizontal line with lower and upper limits representing "no pain" and "very much pain,"

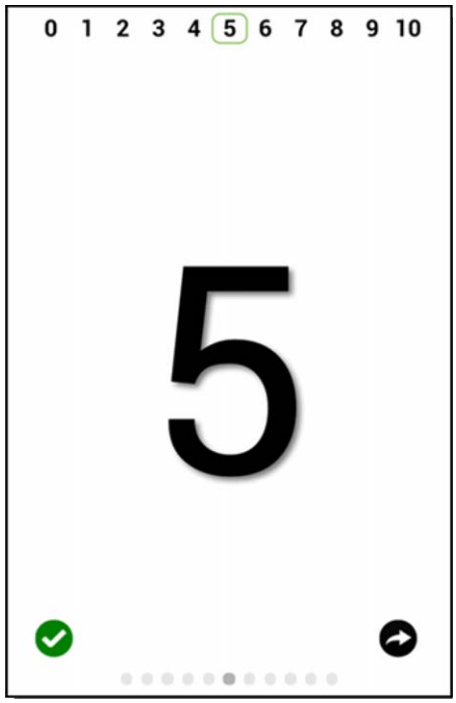

Figure 1 Screenshot of the eNRS-11. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

respectively. We used a mechanical version (mVAS) which consisted of a marker that the children had to move along the line to the point corresponding to the maximum intensity of their most frequent pain during the last 3 months. The pain intensity score was obtained by measuring the distance between the "no pain" limit to the point selected by the marker. The VAS has been shown to provide reliable and valid data when used with children over the age of 5 [28]. However, it is only recommended for use with children above the age of 8 [15]. The procedure used with the electronic version of the VAS (eVAS) was the same as the one used with the mechanical version. Figure 3 shows a screenshot of the eVAS.

The Colored Analogue Scale

CAS consists of a 100 mm-long triangular shape. The anchors are "no pain" and "most pain." However, in this study, we used "very much pain" as the upper limit so that all the scales were homogeneous and there were no potential anchor effects. This procedure was also used in a previous study [23] with good results. At the bottom, CAS is 10 mm wide and at the top it is 30 mm. CAS is also color graduated from white (lower anchor) to dark red (upper anchor). The procedure was the same as with the mVAS: participants had to move a marker along the triangle to the maximum intensity of their most frequent pain in the last three months. CAS

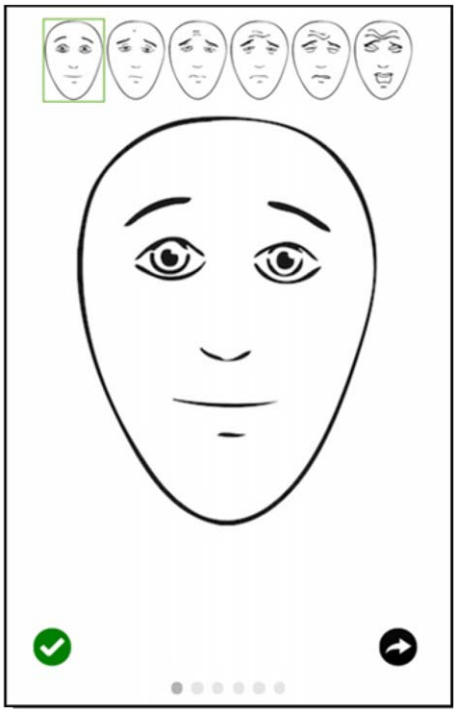

Figure 2 Screenshot of the eFPS-R.
Note: Used and modified with permission from IASP (see www.iasp-pain.org/FSPR). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

has been shown to have good psychometric properties when used in children and adolescents between 5- and 17-year old [28–30]. The electronic version of the CAS (eCAS) works just like the traditional version. Figure 4 shows a screenshot of the eCAS.

Fatigue/Tiredness

Fatigue/Tiredness was measured using a numerical rating scale where 0 corresponded to "no fatigue/tiredness" and 10 to "very much fatigue/tiredness."

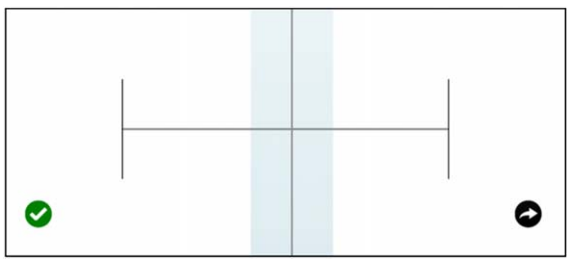

Figure 3 Screenshot of the eVAS. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

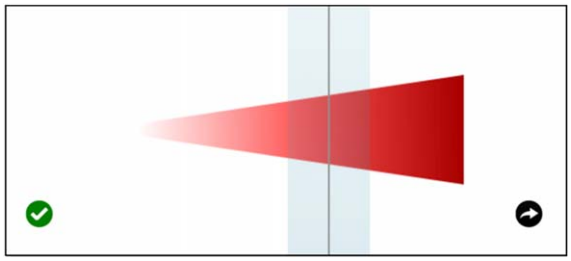

Figure 4 Screenshot of the eCAS. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Catastrophizing

Catastrophizing was measured using the child version of the Pain Catastrophizing Scale (PCS–C) [31]. PCS-C is a 13-item instrument that evaluates catastrophic thinking about pain. On a 5-point Likert scale from 0 (not at all) to 4 (always), children have to indicate the degree to which they agree with each of the statements. The PCS-C has shown good psychometric properties when used in children and adolescents between 8- and 16-year old [31]. In this study we used the Catalan version of the PCS-C [32]. This version was internally consistent (Cronbach's Alpha = 0.86) in our sample.

Procedure

This study was approved by the Department of Education of the Catalan Government and by the participating schools. Participants were recruited from two public schools, which were chosen for their proximity. It is, therefore, a convenience sample.

A letter explaining the study was sent to parents of all the children enrolled in grades 7 to 12. When the informed consent signed by the parents had been returned, an individual interview with each participant was conducted during school hours by two experienced pediatric pain researchers.

First, participants were asked to remember whether they had experienced pain in the previous three months (the question used was "have you had pain in any area of your body in the last 3 months?") and if the answer was positive they were asked to detail the location of their most frequent pain. Next, they had to evaluate the maximum intensity of this pain with the four intensity pain scales: half of the sample did this with the traditional version of the scales and the other half with the electronic version. Then, participants were asked about their level of fatigue and completed the PCS-C, and some other questionnaires. Information on these other questionnaires is not provided here as it is of no interest for the objectives of this study. They were used as a distracting task. This task took about 30 min. After they had completed these questionnaires, participants were

asked again to report the maximum intensity of their most frequent pain. This time, participants who first reported their pain intensity on the traditional versions used the electronic versions and vice versa. Finally, participants reported whether they preferred the traditional or the electronic versions of the scales. The overall interview took about 50 min. The Ipods used in this study were the property of the research team. Each interviewer had an Ipod that was lent to the participant to respond to the questions during the individual interview.

Data Analysis

Data analyses were conducted using SPSS v.17.0 and MedCalc v.12.4.0 for Windows. First, means and standard deviations of pain intensity scores were calculated for both versions of each scale (traditional vs electronic). Then, *t*-tests were used to ensure that the order in which the scales had been presented had had no effect on the pain intensity scores.

To test the one-dimensionality of the scales the FACTOR version 8.02 [33] and the Mplus version 5.1 programs were used [34].

Validity

Construct validity was determined by 1) conducting a confirmatory factor analysis (CFA) of the electronic versions of the four scales and 2) by evaluating convergent and discriminant validity.

Before conducting the CFA, we first obtained the measures of sampling adequacy (Kaiser-Meyer-Olkin [KMO]) and Barlett's test of sphericity from the FACTOR program. Then, we used Mplus to test whether the four electronic scales measured a common factor. We used the mean-adjusted maximum likelihood (MLM) as a factor extraction method. To analyze the adjustment, we used the following fit indices: the root mean square error of approximation (RMSEA), for which a value of less than 0.08 shows a moderate fit and a value of less than 0.05 indicates a good fit [35]; the standardized root mean square residual (SRMR) [36], for which a value of less than 0.08 shows a good fit [37]; the comparative fit index (CFI) [38]; and the Tucker–Lewis index (TLI) [39], for which a value greater than 0.90 indicates a good fit.

Convergent validity was calculated by correlating participants' scores on the electronic versions of the scales with those on the corresponding scales in their traditional versions. It was assumed that correlation among these scales would be high because all the measures assess the same construct [23]. Discriminant validity was estimated by comparing the magnitude of the correlation coefficients between ratings on the four scales in their electronic versions with the magnitude of the correlation coefficients between ratings on the four scales and the scores on the NRS-Fatigue. It was assumed that correlations between the pain intensity scales would be significantly higher than those between these scales and the fatigue scores.

Criterion validity was also analyzed. As the measures were taken at the same time, concurrent validity was calculated by correlating participants' pain intensity ratings on the four scales with participants' scores on the PCS-C. It was assumed that catastrophizing can be predicted from pain intensity.

Agreement

To determine whether reports provided by electronic versions of the scales are concordant with those provided by the traditional ones we used the Bland–Altman method [40]. This method has been successfully used to determine agreement between reports from different pediatric pain intensity scales in previous studies [8,23,41]. The Bland–Altman method uses a graph to represent the difference between the scores reported by each subject with two different scales or methods and compares it with the average of these two scores. We conducted analyses to determine the agreement between the scores reported on vNRS-11–eNRS-11, FPS-R–eFPS-R, mVAS–eVAS, and CAS–eCAS

To interpret the data of the Bland–Altman method we set a priori a maximum limit of agreement to each scale. This value corresponded to the minimal clinically significant difference (MCSD) for children and adolescents when using each of the pain intensity scales. Thus, the maximum limit of agreement for the comparison between vNRS-11 and eNRS-11 was ± 1 [19,42], for FPS-R and eFPS-R it was ± 2 [43,44], for mVAS and eVAS it was -16–$+18$ (MCSD when children were asked with a 30 min interval and the pain had not changed) [45], and for CAS and eCAS it was ± 20 [46].

A previous study on the agreement between the reports provided by the traditional versions of the four scales [23] elicited a written reflection about the suitability of using a 95% confidence interval (CI) for evaluating subjective variables like pain intensity in pediatric populations [47]. The authors stated that it may be more appropriate to use an 80% CI instead of the 95% CI established by Bland and Altman [40]. For this reason, in this study we provide data for both 80% and 95% CI.

Preference

Finally, we conducted a *z*-test to identify which version of these scales (traditional vs electronic) was preferred.

Table 2 Presentation order effects

	Mean of NRS-11	Mean of CAS	Mean of FPS-R	Mean of mVAS
Traditional version first	6.64	64.85	5.61	59.29
Traditional version second	6.79	69.90	5.68	64.52
t-test t(*P*)	0.60 $P = 0.55$	2.14 $P = 0.03$*	0.25 $P = 0.81$	1.92 $P = 0.06$
Electronic version first	6.87	67.61	5.61	64.57
Electronic version second	6.55	64.95	5.49	59.05
t-test	1.3 $P = 0.20$	1.17 $P = 0.25$	0.41 $P = 0.68$	2.2 $P = 0.03$*

* $P < 0.05$

Results

Presentation Order Effects

Ninety-eight participants first used the traditional versions of the scales to report their pain intensity, whereas 82 used them second.

Generally speaking, the presentation order of the scales (traditional vs electronic versions and vice versa) had no influence on the participants' scores except when the CAS was administered first in its traditional form and when the VAS was presented first electronically. In these cases, effect size power was calculated and proved to be small in both cases (Cohen's d= 0.32; $r = 0.16$) [48]. Table 2 shows all the t-tests conducted.

Construct Validity

Confirmatory Factor Analysis

The KMO sampling adequacy index (0.82) and Barlett's test of sphericity (634.7; df = 6; $P < 0.00001$) showed good fit and proper sample correlation between items, indicating that our data are appropriate for factorization. With a χ^2 of 1.576 and 2 degrees of freedom, we accepted the null hypothesis that all electronic scales measure one common factor. The goodness of fit index showed an excellent fit (CFI = 1.00 and TLI = 1.00; SRMR = 0.008; RMSEA < 0.0001). Table 3 shows the loadings of each scale on the factor.

Table 3 Unrotated loading matrix

Rating scale	Factor 1	Communality
eNRS-11	0.87	0.76
eCAS	0.93	0.86
eFPS-R	0.67	0.45
eVAS	0.97	0.94

eNRS-11 = electronic version of Numerical Rating Scale; eCAS = electronic version of Colored Analogue Scale; eFPS-R = electronic version of Faces Pain Scale - Revised; eVAS = electronic version of visual analogue scale.

Convergent and Discriminant Validity

Convergent validity was supported because the correlation coefficients between the electronic and traditional versions of the four scales are all higher than 0.5 (between 0.58 and 0.86; see Table 4). The correlations between the reports provided by each electronic version of the scales and its traditional counterpart are also high (between 0.88 and 0.92; see Table 4).

Discriminant validity was also supported because the magnitude of the correlation between the four electronic versions of the scales with each other was greater than the correlations between these scales and the NRS-Fatigue (see Table 4).

Criterion Validity

Concurrent validity was supported by moderate and positive correlations between the reports provided by the four electronic scales of pain intensity and those provided by the PCS-C (between 0.29 and 0.32; see Table 4).

Agreement

Data about agreement are shown in Table 5 and graphs can be found in Appendix.

The agreement between the reports provided by the traditional and electronic versions of all the scales has been demonstrated for the 80% CI. If we take into account the results at 95% CI, then we find an agreement between reports provided by the two versions of the FPS-R, the VAS and the CAS. The only scale that did not show an agreement between the reports of its two versions at 95% CI was the NRS-11.

Preference

Eleven participants were excluded from this analysis because of missing data. Of the available participants, 134 (79%) preferred to report their pain using the electronic version of the scales, whereas 35 (21%) preferred the traditional versions. The difference was statistically significant ($z = 7.53$, $P < 0.001$). Preference was not

Table 4 Validity of the four electronic versions of pain intensity scales

Construct validity				
Convergent validity				
	eNRS-11	eCAS	eFPS-R	eVAS
vNRS-11	0.90***	0.75***	0.60***	0.81***
CAS	0.77***	0.88***	0.61***	0.86***
FPS-R	0.58***	0.61***	0.89***	0.62***
mVAS	0.79***	085***	0.65***	0.92***
Discriminant validity				
eNRS-11-eCAS – eNRS-11-NRS-Fatigue			$z = 7.1$***	
eNRS-11-eFPS-R – eNRS-11-NRS-Fatigue			$z = 2.4$*	
eNRS-11-eVAS – eNRS-11-NRS-Fatigue			$z = 8.08$***	
eFPS-R-eNRS-11 – eFPS-R-NRS-Fatigue			$z = 5.07$***	
eFPS-R-eCAS – eFPS-R-NRS-Fatigue			$z = 4.91$***	
eFPS-R-eVAS – eFPS-R-NRS-Fatigue			$z = 5.80$***	
eCAS-eNRS-11 – eCAS-NRS-Fatigue			$z = 8.37$***	
eCAS-eFPS-R – eCAS-NRS-Fatigue			$z = 3.22$**	
eCAS-eVAS – eCAS-NRS-Fatigue			$z = 12.99$***	
eVAS-eNRS-11 – eVAS-NRS-Fatigue			$z = 8.06$***	
eVAS-eFPS-R – eVAS-NRS-Fatigue			$z = 3.01$**	
eVAS-eCAS – eVAS-NRS-Fatigue			$z = 11.43$***	
Criterion validity (*concurrent*)				
eNRS-11 – PCS-C^n			$r = 0.30$***	
eCAS – PCS-C^n			$r = 0.32$***	
eFPS-R – PCS-C^n			$r = 0.29$***	
eVAS – PCS-C^n			$r = 0.31$***	

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

n = results based on information from 159 participants

NRS-11 = Numerical Rating Scale; CAS = Colored Analogue Scale; FPS-R = Faces Pain Scale - Revised; mVAS = mechanical version of visual analogue scale.

influenced by the order of presentation of the scales (see Table 6).

Discussion

The findings from this study provide important information regarding the validity of the reports provided by the electronic versions of four scales widely used to measure pediatric pain intensity: the Numerical Rating Scale (NRS-11), the Faces Pain Scale-Revised (FPS-R), the mechanical visual analogue scale (mVAS), and the Colored Analogue Scale (CAS). The results also provide evidence on the agreement between the scores reported by traditional and electronic versions of the scales that are included in *Painometer* [17], a mobile application developed to help assess people with pain. Pain intensity reports provided by all scales proved to be valid. The construct validity was supported via 1) strong loadings on a factor score representing global pain intensity, 2) significant associations between reports provided by electronic and traditional versions of the scales (convergent validity), and [3]) higher correlations between reports provided by the four electronic versions of the scales with each other than correlations between

Table 5 Limits of agreement according to the Bland–Altman method

			Limits of Agreement			
			95% CI		80% CI	
Scale	Mean Difference	Maximum Limit of Agreement	Lower	Upper	Lower	Upper
eNRS-11–vNRS-11	−0.01	±1	−1.5	1.5	−0.97	0.95
eFPS-R–FPS-R	−0.13	±2	−1.9	1.6	−1.27	1.0
eVAS–mVAS	−0.9	−16–+18	−15.1	13.3	−10.2	8.4
eCAS–CAS	−1.1	±20	−15.9	13.7	−10.8	8.5

Table 6 Preference according to order presentation

	Traditional–Electronic (90)	Electronic–Traditional (79)	*z* (*P*)
Traditional version	21 (23%)	14 (18%)	0.36 (>0.05)
Electronic version	69 (77%)	65 (82%)	0.80 (>0.05)

reports from these scales and those from the NRS-Fatigue (discriminant validity). Concurrent validity was supported via moderate positive correlations between pain intensity reports provided by the four electronic versions of the scales, and scores on the PCS-C.

Reports provided with the traditional pain intensity scales were found to be in agreement with those reported by their electronic counterparts when they were analyzed at an 80% CI, suggesting, as hypothesized, that these scales could be used interchangeably. Our results are in line with those reported by Castarlenas et al. [8] who found an agreement between reports provided by the verbally administered NRS-11 and its electronic version.

There are certain limitations that should be kept in mind when interpreting the study's findings. First, although the number of participants is appropriate for the purposes of the study, the convenience sample consisted largely of essentially healthy schoolchildren and they may or may not be representative of the population. Thus, replication of the current findings in other samples of young people with acute or chronic pain problems would help to establish their reliability. Second, although we counterbalanced the presentation of the scales (traditional vs electronic version), we did not counterbalance the presentation order of the scales. Although previous studies have found that the order in which scales are presented does not influence the ratings obtained [21], it is possible that order effects might have influenced in some unknown ways. Third, although one advantage of using electronic versions of pain intensity self-reports in mobile devices is the assessment of pain in real time, in this study we asked participants to recall their maximum pain intensity in the last three months. This procedure was implemented because of the characteristics of our sample: not all the participants were experiencing pain at the time of the interview; however, this procedure has been successfully used in previous studies [e.g., 8,20,23]. Fourth, the time elapsed between the two assessments was 30 min, which may not be long enough to prevent memory effects. Nevertheless, participants performed a distracting task in between, they did not know that we were going to ask them for their pain intensity a second time, and they had not seen the values of the previous assessments.

However, in spite of these limitations, this study advances our knowledge by giving specific additional information on the validity and the concordance of the reports obtained with electronic versions of four of the most widely used scales to assess pediatric pain intensity. Furthermore, the results of our study demonstrate that pain intensity reports provided by *Painometer* are valid and that there is an agreement between the reports provided by traditional and electronic versions of the 0–10 Numerical Rating Scale, the mechanical visual analogue scale, the Faces Pain Scale—Revised and the Colored Analogue Scale at 80% and 95% CI, with the exception that the two versions of the NRS-11 only agree at 80% CI. These results have important implications for researchers and clinicians working with pediatric populations with pain. First, pain reports provided by the electronic versions of the scales are valid, and young people also seem to prefer the electronic versions of the scales to the traditional ones, at least in general (we did not ask participants about each individual scale). Second, researchers and clinicians can compare pain intensity scores reported with a traditional form of one of these scales (e.g., during a face-to-face interview) with pain intensity scores obtained from the same patient with the electronic version of the same scale (for example during telephone follow-ups). And, third, *Painometer* is a mobile application that can help to assess people with pain. These results are in line with those of recent studies reporting on the agreement of the electronic version of the NRS-11 [8] and the usability of Painometer [17].

Further work is needed to analyze other psychometric characteristics of the pain intensity reports provided by electronic versions of the scales: reliability, sensibility, etc. However, on the basis of these findings, the electronic versions of the scales, at least as they are presented in the mobile application *Painometer*, can be recommended for assessing pain intensity in young people.

Acknowledgments

The authors would like to thank the participating schools (Escola Puigcerver and Col·legi Cardenal Vidal i Barraquer) for their interest and collaboration in the study. This study was partly funded by the Government of Catalonia, the Spanish Ministry of Science (grant # PSI2012-32471), *RecerCaixa*, Universitat Rovira i Virgili (PFR program), and the *Fundació La Marató de TV3*. ES-R is supported by a doctoral grant from the Catalan Government. RV is supported by a postdoctoral grant from the Catalan Government. JM's work is supported by the *Institució Catalana de Recerca i Estudis Avançats* (ICREA-Acadèmia).

References

1 Stinson JN, Jibb L. New information and communication technologies for pain. In: McGrath PJ, Stevens BJ, Walker SM, Zempsky WT, eds. Oxford

Textbook of Pediatric Pain. Oxford: University Press; 2013:569–77.

2 Hicks CL, von Baeyer CL, McGrath PJ. Online psychological treatment for pediatric recurrent pain: A randomized evaluation. J Pediatr Psychol 2006; 31(7):724–36.

3 Stinson JN, Mcgrath PJ, Hodnett ED, et al. An internet-based self-management program with telephone support for adolescents with arthritis: A pilot randomized controlled trial. J Rheumatol 2010;37: 1944–52.

4 Palermo TM, Valenzuela D, Stork PP. A randomized trial of electronic versus paper pain diaries in children: Impact on compliance, accuracy, and acceptability. Pain 2004;107:213–9.

5 Stinson JN, Petroz GC, Tait G, et al. e-Ouch: Usability testing of an electronic chronic pain diary for adolescents with arthritis. Clin J Pain 2006;22: 295–305.

6 Stinson JN, Petroz GC, Stevens, BJ, et al. Working out the kinks: Testing the feasibility of an electronic pain diary for adolescents with arthritis. Pain Res Manag 2008;13:375–82.

7 Wood C, von Baeyer CL, Falinower S, et al. Electronic and paper versions of a faces pain intensity scale: Concordance and preference in hospitalized children. BMC Pediatr 2011;11:87

8 Castarlenas E, Sánchez-Rodríguez E, de la Vega R, Roset R. Miró J. Agreement between verbal and electronic versions of the numerical rating scale (NRS-11) when used to assess pain intensity in adolescents. Clin J Pain 2015;31:229–34.

9 Jibb L, Stinson JN, Nathan P, et al. Pain squad: Usability testing of a multidimensional electronic pain diary for adolescents with cancer. J Pain 2012, 13, 4 Supplement, S23.

10 Mcgrath PJ, Watters C, Moon E. Technology in pediatric pain management. In: Finley G, Mcgrath PJ, Chambers CT, eds. Bringing Pain Relief to Children. Totowa: Humana Press; 2006:159–76.

11 Stinson JN. Improving the assessment of pediatric chronic pain: Harnessing the potential of electronic diaries. Pain Res Manag 2009;14:59–64.

12 Lewandowski AS, Palermo TM, Kirchner HL, Drotar D. Comparing diary and retrospective reports of pain and activity restriction in children and adolescents with chronic pain conditions. Clin J Pain 2009;25:299–306.

13 Morren M, van Dulmen S, Ouwerkerk J, Bensing J. Compliance with momentary pain measurement using electronic diaries: A systematic review. Eur J Pain 2009;13:354–65.

14 Walker L, Sorrells S. Brief report: Assessment of children's gastrointestinal symptoms for clinical trials. J Pediatr Psychol 2002;27:303–7.

15 McGrath PJ, Walco GA, Turk DC, et al. Core outcome domains and measures for pediatric acute and chronic/recurrent pain clinical trials: PedIMMPACT recommendations. J Pain 2008;9(9):771–83.

16 Stanford EA, Chambers CT, Craig KD. The role of developmental factors in predicting young children's use of a self-report scale for pain. Pain 2006;120: 16–23.

17 de la Vega R, Roset R, Castarlenas E, et al. Development and testing of Painometer: A Smartphone app to assess pain intensity. J Pain 2014;15:1001–7.

18 Stinson JN, Stevens BJ, Feldman, BM, et al. Construct validity of a multidimensional electronic pain diary for adolescents with arthritis. Pain 2008;136: 281–92.

19 Bailey B, Daoust R, Doyon-Trottier E, Dauphin-Pierre S, Gravel J. Validation and properties of the verbal numeric scale in children with acute pain. Pain 2010;149:216–21.

20 Castarlenas E, Miró J, Sánchez-Rodríguez E. Is the verbal numerical rating scale a valid tool for assessing pain intensity in children below 8 years of age? J Pain 2013;14:297–304.

21 Miró J, Castarlenas E, Huguet A. Evidence for the use of a numerical rating scale to assess the intensity of pediatric pain. Eur J Pain 2009;13: 1089–95.

22 Pagé MG, Katz J, Stinson JN, et al. Validation of the numerical rating scale for pain intensity and unpleasantness in pediatric acute postoperative pain: Sensitivity to change over time. J Pain 2012; 13:359–69.

23 Sánchez-Rodríguez E, Miró J, Castarlenas E. A comparison of four self-report scales of pain intensity in 6- to 8-year-old children. Pain 2012;153: 1715–9.

24 von Baeyer CL, Spagrud LJ, McCormick JC, et al. Three new datasets supporting use of the Numerical Rating Scale (NRS-11) for children's self-reports of pain intensity. Pain 2009;143:223–7.

25 Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The faces pain scale-revised: Toward a common metric in pediatric pain measurement. Pain 2001;93(2):173–83.

26 Miró J, Huguet A. Evaluation of reliability, validity, and preference for a pediatric pain intensity scale: The Catalan version of the faces pain scale–revised. Pain 2004;111:59–64.

27 Stinson JN, Kavanagh T, Yamada J, Gill N, Stevens B. 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 2006; 125(1–2):143–57.

28 McGrath PA, Seifert CE, Speechley KN, et al. A new analogue scale for assessing children's pain: An initial validation study. Pain 1996;64:435–43.

29 McGrath PA, Gillespie J. Pain assessment in children and adolescents. In: Turk DC, Melzack R, eds. Handbook Assessment. New York: Guilford Press; 2001:97–118.

30 Miró J, Huguet A, Nieto R, Paredes S. Baos J, The Color Analog Scale: A valid instrument for measuring pain intensity in Catalan-speaking children? Rev Esp Anestesiol Reanim 2007;54:208–12.

31 Crombez G, Bijttebier P, Eccleston C, et al. The child version of the pain catastrophizing scale (PCS-C): A preliminary validation. Pain 2003;104(3):639–46.

32 Miró J, Huguet A, Nieto R. The child version of the pain catastrophizing scale (PCS-C): A sueful tool to predict pain intensity and pain-related disability in different pain contexts 2008. Poster presented at 12th World Congress on Pain. G. lasgow, Scotland.

33 Lorenzo-Seva U, Ferrando PJ. FACTOR: A computer program to fit the exploratory factor analysis model. Behav Res Methods, Instruments Comput 2006;33:88–91.

34 Muthén LK, Muthén BO. Mplus User's Guide: Statistical Analysis with Latent Variables, 4th edition. Los Angeles, CA: Muthén & Muthén; 2007.

35 Steiger J. Structural model evaluation and modification: An interval estimation approach. Multivariate Behav Res 1990;25:173–80.

36 Bentler PM. EQS Structural Equations Program Manual. Encino, CA: Multivariate Software; 1995.

37 Hu L-T, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 1999;6:1–55.

38 Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107:238–46.

39 Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika 1973; 38:1–10.

40 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;1:307–10.

41 Bailey B, Bergeron S, Gravel J, Daoust R. Comparison of four pain scales in children with acute abdominal pain in a pediatric emergency department. Ann Emerg Med 2007;50(4):379–83, 383.e1–2.

42 Voepel-Lewis T, Burke C, Jeffreys N, Malviya S, Tait A. Do 0-10 numeric rating scores translate into clinically meaningful pain measures for children?. Anesth Analg 2011;112:415–21.

43 Tomlinson D, von Baeyer CL, Stinson JN, Sung L. A systematic review of faces scales for the self-report of pain intensity in children. Pediatrics 2010;126: e1168–98.

44 vonBaeyer CL. Children's self-reports of pain intensity: Scale selection, limitations and interpretation. Pain Res Manag 2006;11:157–62.

45 Bailey B, Gravel J, Daoust R. Reliability of the visual analog scale in children with acute pain in the emergency department. Pain 2012;153(4):839–42.

46 Bulloch B, Tenenbein M. Assessment of clinically significant changes in acute pain in children. Acad Emerg Med 2002;9:199–202.

47 von Baeyer CL. Reported lack of agreement between self-report pain scores in children may be due to a too strict criterion for agreement. Pain 2012;153:2152–4.

48 Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition. Hillsdale: Lawrence Earlbaum; 1988.

Appendix

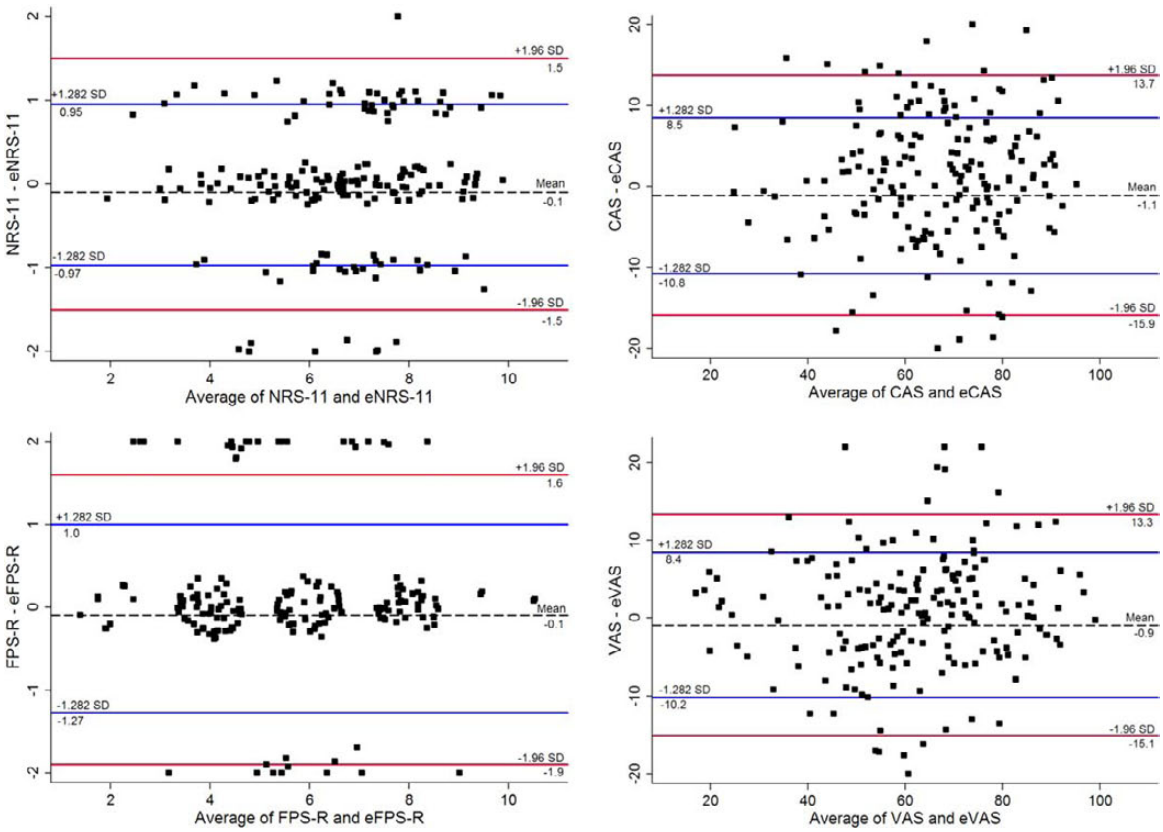

Figure A1 Bland & Altman graphs: agreement between traditional vs electronic scales. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]