

## Concurrent validity and interobserver reliability of the EPInfant pediatric perceived exertion rating scale among healthy Chilean children

Iván Rodríguez-Núñez, Kinesiologist<sup>a,b,c</sup> and Carlos Manterola, M.D.<sup>c,d,e</sup>

### ABSTRACT

The objective of this study was to determine the validity and reliability of the EPInfant pediatric perceived exertion rating scale among Chilean children. Children aged 8-12 years old were selected; their heart rate, perceived exertion and workload were recorded during two consecutive Chester tests performed with a one week interval. Pearson's r coefficient and the intraclass correlation coefficient were estimated. Thirty-one children (17 boys) were included in the study. A strong correlation was observed between perceived exertion and reference criteria, in addition to a high intraclass correlation coefficient. The average difference between the perceived exertion in each test was 1.4 (95% CI: -1.4 to 4.1) and 0.2 (95% CI: -1.8 to 2.2) for boys and girls, respectively. The EPInfant scale showed an acceptable criterion validity and interobserver reliability among children during an incremental step test.

**Key words:** reliability and validity, children, scales, exercise, exertion.

<http://dx.doi.org/10.5546/aap.2016.eng.343>

### INTRODUCTION

Perceived exertion (PE) rating scales are used to measure the feeling caused by metabolic changes during exercise. Thus, the Borg scale has been widely used and has demonstrated acceptable validity levels among adults.<sup>1</sup> However, there is no favorable evidence regarding its performance in the pediatric population.<sup>2,3</sup>

- a. School of Kinesiology, School of Health Sciences, Universidad San Sebastián, Concepción, Chile.
- b. Kinesiology Degree, School of Health, Universidad San Sebastián, Concepción, Chile.
- c. Doctoral Program in Medical Sciences, School of Medicine, Universidad de la Frontera, Temuco, Chile.
- d. Department of Surgery, Universidad de La Frontera, Chile.
- e. Biomedical Sciences Research Center, Universidad Autónoma de Chile.

*E-mail address:* Iván Rodríguez Núñez, Kinesiologist: [ivan.rodriguez@uss.cl](mailto:ivan.rodriguez@uss.cl)

*Funding:* This study was funded by the School of Kinesiology, School of Health Sciences, Universidad San Sebastián, Concepción, Chile.

*Conflict of interest:* None.

Received: 9-27-2015

Accepted: 1-27-2016

The psychometric properties of these scales have been recently described and it has been observed that, to a great extent, they depend on the subjects' cognitive development and the type of instrument used,<sup>4</sup> thus the pediatric perceived exertion rating scale (EPInfant) was developed. The EPInfant scale uses numerical descriptors (0 to 10), verbal descriptors in Spanish, and a set of pictures that illustrate a child running at progressively increasing intensities.<sup>5</sup> It has evidenced a high validity level among 13-to-15-year-old adolescents,<sup>5,6</sup> but it has been scarcely assessed among younger children.<sup>7</sup>

Therefore, the objective of this study was to establish the concurrent criterion validity and interobserver reliability of the EPInfant scale among 8-to-12-year-old healthy children during an incremental exercise test.

### METHODS

This study implemented a test-retest design and included children aged between 8 and 12 years old from a school located in Concepción, Chile. Children with obesity, cognitive impairment, cardiorespiratory or motor neurone disease were excluded. The sample size was estimated based on an intraclass correlation coefficient (ICC) of 0.7, a 5% type I error, a 10% type II error, and a confidence interval (CI) for ICC below 0.15. Therefore, the minimum sample needed for this study was 14 boys and 14 girls.

*Study protocol:* One week before starting with the tests, children's age, sex and anthropometric outcome measures were recorded. Weight and height were measured using an analog scale and a tape measure. Also during this visit, participants became familiar with the EPInfant scale, as per Rodríguez et al.<sup>5</sup>

The exercise test used in this study was the Chester test.<sup>8</sup> The test was done twice, with a one week interval. The Chester test consists in going up and down a 20 cm step at an increasing pace set by a signal sound. It includes five levels, and each level is relative to the number of times the subject goes up and down the step with both feet in one minute. Each level lasts two minutes, for

a total of 10 minutes.

Before completing the Chester test, children rested for 10 minutes, their heart rate (HR) was recorded (Polar® monitor) and the maximum HR was estimated using Tanaka's formula:  $208 - (0.7 \times \text{age})$ .

During the Chester test, HR and PE were recorded by two separate observers during the last 20 seconds of each minute. In addition, PE for tests 1 and 2 was measured by two different observers. Workload (watts) was estimated based on the following formula:  $\text{step height (m)} \times \text{number of cycles} \times \text{weight (kg)} \times 0.16357$ .<sup>9</sup>

This study was approved by the Ethics Committee of Universidad San Sebastián; each participant's legal guardian was asked to sign an informed consent. The study was conducted between September and December 2014.

**Statistical analysis:** The MedCalc 14.12 statistical software was used to establish descriptive statistics with average and standard deviation (SD) estimations, and normality was verified using the Shapiro-Wilk test. A repeated measures analysis of variance (ANOVA) with Bonferroni's correction was used to assess differences in HR, workload and PE among the Chester test's different intensity levels. Student's t test was used to assess differences between genders for independent samples.

The concurrent criterion validity analysis was done using linear regression between PE and HR, and between PE and workload. The ICC was estimated to establish interobserver reliability.

In addition, average PE differences and their limits of agreement were estimated using the Bland-Altman method. Validity and reliability results were described as absolute value and 95% CI. A value of  $p < 0.05$  was considered significant.

## RESULTS

The study included 31 children (17 boys) aged between 8 and 12 years old. The sample characteristics are reported in *Table 1*.

Exercise intensity during the Chester test was significantly higher among boys versus girls. Besides, HR and PE increased significantly with each intensity level ( $p < 0.05$ ). No differences in the maximum levels of HR and PE were observed between boys and girls.

Only among girls, a lower HR was observed in the first Chester test level, but no differences were established in the other levels. In addition, and only among boys, a lower PE level was observed in every intensity level during the first Chester test.

In relation to criterion validity, a strong positive correlation was observed between PE and HR, and between PE and workload (*Figure 1. A and B*). A significant difference between boys and girls was observed only in relation to the regression slope between PE and workload ( $p = 0.03328$ ). Finally, all regression models were statistically significant ( $p < 0.0001$ ). *Table 2* shows the components of regression equations.

In terms of reliability, the general ICC was 0.88 (95% CI: 0.82-0.92) and 0.94 (95% CI: 0.91-0.96) for boys and girls, respectively. In each Chester test level, reliability was moderate to strong, and it reached statistical significance among both boys and girls, except for level I (*Table 3*). Besides, the mean difference between PE measurements was 1.4 (95% CI: 4.1-1.4) and 0.2 (95% CI: 2.2-1.8) for boys and girls, respectively (*Figure 1. C and D*).

## DISCUSSION

The main findings of this study confirm that the EPInfant scale has an acceptable validity and reliability among 8-to-12-year-old children during the step test.

These findings are consistent with validity studies conducted using other scales in the Anglo-Saxon population and in other not Spanish-speaking populations, which have demonstrated a high validity level during incremental exercise tests in children (Pearson's  $r > 0.85$ ).<sup>10, 11</sup> In our study, the coefficient of determination ( $r^2$ ) was higher than 0.9 in both boys and girls, which demonstrates a high interdependence between reference criteria and PE (EPInfant) during exercise.

In addition, the reliability of these instruments has been scarcely verified in children. Leung et al. observed an ICC above 0.85 among children

TABLE 1. General characteristics of the study sample

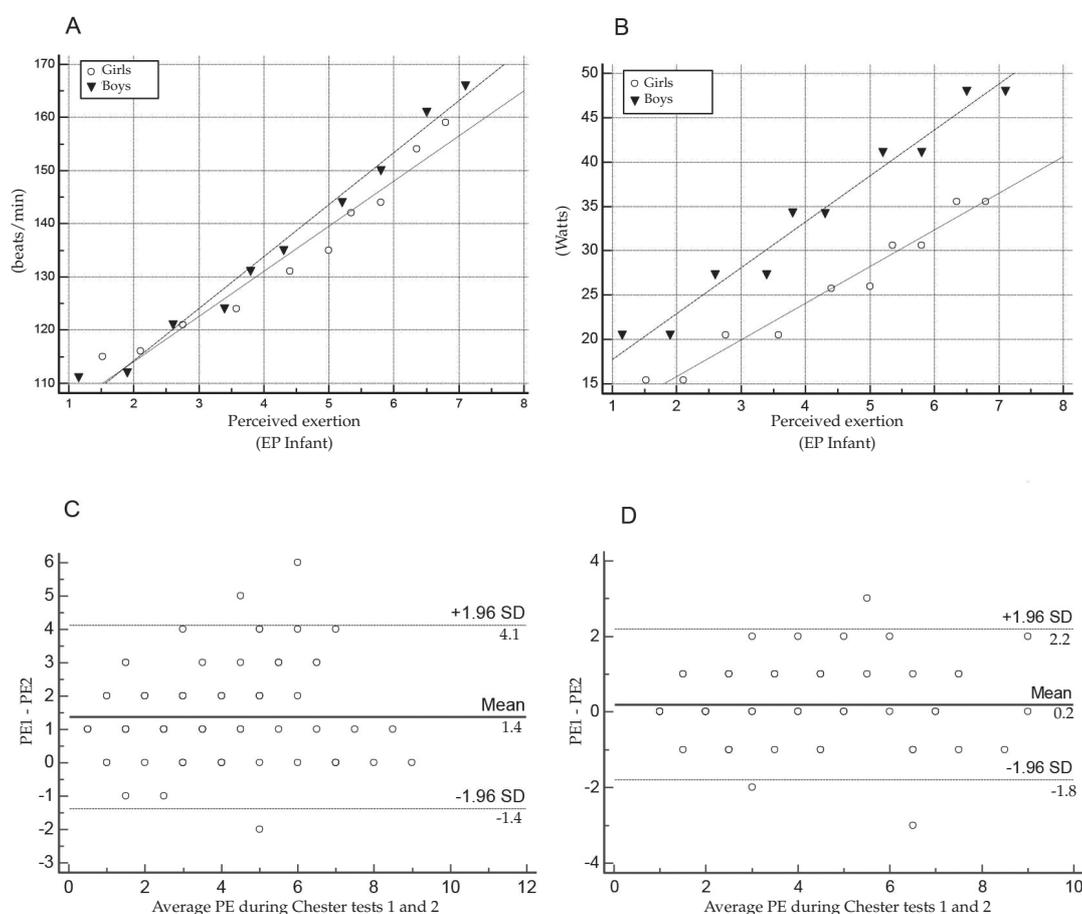
	Boys (n= 17)	Girls (n= 14)
Age (years)	10.1 ± 0.6	8.7 ± 0.9
Weight (kg)	42.1 ± 9.4	31.4 ± 5.9
Height (cm)	144.9 ± 11.0	133.5 ± 7.4
Theoretical maximum HR (beats/min)	201.0 ± 0.8	201.5 ± 0.8
HR at the end of Chester test 1	160.6 ± 15.6	155.5 ± 14.6
HR at the end of Chester test 2	164.2 ± 13.8	162.1 ± 17.9
Workload at the end of the Chester test (watts)	47.9 ± 10.3	35.5 ± 6.1

HR: heart rate; Chester test: Chester incremental step test. Results are described as average and standard deviation.

aged 10 and 11 years old during an incremental exercise test.<sup>2</sup> Pfeiffer et al. noted an acceptable reliability level for the measurement of PE during exercise, especially in scales that have been adapted to the pediatric population (OMNI scale).<sup>3</sup>

It is worth considering that ICC accounts for the proportion of variability not caused by a measurement error.<sup>12</sup> Thus, given the high ICC observed in both boys and girls, we may confirm that the EPInfant scale has a low risk for a random measurement error. In addition, in the reliability

FIGURE 1. Criterion validity and reliability of the EPInfant scale. A shows the correlation between heart rate and perceived exertion during the Chester test. B shows the correlation between workload and perceived exertion during the Chester test. C is a Bland-Altman plot showing individual differences among boys. D is a Bland-Altman plot showing individual differences among girls.



PE: perceived exertion.  
Results are described as average and 95% confidence interval.

TABLE 2. Linear regression analysis for the EPInfant pediatric perceived exertion rating scale described as heart rate function (beats/min) and workload during the Chester test (watts)

Outcome	Subgroup	Slope	Intercept	r	r <sup>2</sup>	P value
Heart rate	All	9.0	96.26	0.97	0.95	< 0.0001
	Boys	9.7	94.9	0.99	0.98	< 0.0001
	Girls	8.3	97.7	0.97	0.95	< 0.0001
Workload	All	4.4	10.7	0.84	0.71	< 0.0001
	Boys	5.0	12.9	0.98	0.96	< 0.0001
	Girls	4.0	7.7	0.98	0.96	< 0.0001

TABLE 3. Reliability of the EPInfant pediatric perceived exertion rating scale during the Chester test

Boys	Level I	Level II	Level III	Level IV	Level V
PE 1	2.0 ± 1.1	3.4 ± 1.5	4.7 ± 1.9	5.5 ± 1.6	6.5 ± 1.8
PE 2	1.1 ± 0.8*	2.1 ± 1.3*	2.8 ± 1.6*	4.1 ± 1.7*	5.2 ± 1.7*
ICC	0.58	0.67*	0.77*	0.84*	0.73*
Girls	Level I	Level II	Level III	Level IV	Level V
PE 1	1.5 ± 0.6	2.6 ± 1.2	3.8 ± 1.5	5.2 ± 1.6	6.6 ± 2.1
PE 2	1.4 ± 0.6	2.6 ± 1.2	3.6 ± 1.6	4.8 ± 1.6	6.4 ± 1.9
ICC	0.69	0.76*	0.92*	0.89*	0.84*

PE: perceived exertion; Chester test: Chester incremental step test; ICC: intraclass correlation coefficient;

\*  $p < 0.05$ ; results are described as average and standard deviation.

\* A significant difference was observed in the boys' PE between the first and the second assessment (level 1,  $p = 0.002$ ; level 2,  $p = 0.0016$ ; level 3,  $p = 0.0002$ ; level 4,  $p = 0.0003$ ; level 5,  $p = 0.004$ ). Student's t test was used for paired samples.

analysis for each Chester test level, a higher ICC was observed in the three higher intensity levels; therefore, it is likely that physiological stress may have reduced the variability in the reported PE level. This is consistent with some reports suggesting the fundamental role of experience on individuals' perceptual ability regarding exertion.<sup>13</sup>

One of the weaknesses of this study is the difference in age between boys and girls; however, both groups were categorized into Piaget's concrete operational stage. Also, the small number of studied subjects does not allow extrapolating results, so this study should be replicated in other populations.

To sum up, the EPInfant scale demonstrated an adequate criterion validity and interobserver reliability to measure PE in a sample of healthy 8-to-12-year-old children. Future studies are necessary to confirm these results. ■

### Acknowledgments

We would like to thank all children and parents who participated in this study.

### REFERENCES

- Chen MJ, Fan X, Moe ST. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. *J Sports Sci* 2002;20(11):873-99.
- Leung ML, Chung PK, Leung RW. An assessment of the validity and reliability of two perceived exertion rating scales among Hong Kong children. *Percept Mot Skills* 2002;95(3 Pt 2):1047-62.
- Pfeiffer KA, Pivarnik JM, Womack CJ, Reeves MJ, et al. Reliability and validity of the Borg and OMNI rating of perceived exertion scales in adolescent girls. *Med Sci Sports exerc* 2002;34(12):2057-61.
- Rodríguez I, Zambrano L, Manterola C. Validez de criterio de las escalas de medición de esfuerzo percibido en niños sanos: una revisión sistemática y metaanálisis. *Arch Argent Pediatr* 2016;114(2):120-8.
- Rodríguez I, Zenteno D, Cisternas L, Rodríguez P, et al. Construcción y evaluación de EPInfant: una escala para la medición del esfuerzo percibido en población pediátrica. *Arch Argent Pediatr* 2015;113(6):550-7.
- Rodríguez I, Manterola C. Validación inicial de la escala de medición de esfuerzo percibido infantil (EPInfant) en niños chilenos. *Biomédica* 2016;36(1). [Epub 2015 21 Jul].
- Rodríguez Nuñez I. Escala de medición de esfuerzo percibido infantil (EPInfant): validación en niños y adolescentes chilenos. *Rev Chil Pediatr* 2015. [Epub 2015 Oct 8].
- Buckley JP, Sim J, Eston RG, Hession R, et al. Reliability and validity of measures taken during the Chester step test to predict aerobic power and to prescribe aerobic exercise. *Br J Sports Med* 2004;38(2):197-205.
- De Andrade CH, de Camargo AA, de Castro BP, Malaguti C, et al. Comparison of cardiopulmonary responses during 2 incremental step tests in subjects with COPD. *Respir Care* 2012;57(11):1920-6.
- Barkley JE, Roemmich JN. Validity of the CALER and OMNI-bike ratings of perceived exertion. *Med Sci Sports exerc* 2008;40(4):760-6.
- Gros Lambert A, Mahon AD. Perceived exertion: influence of age and cognitive development. *Sports Med* 2006;36(11):911-28.
- Zou GY. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Stat Med* 2012;31(29):3972-81.
- Barroso R, Cardoso RK, do Carmo EC, Tricoli V. Perceived exertion in coaches and young swimmers with different training experience. *Int J Sports Physiol Perform* 2014;9(2):212-6.