DOI: 10.4025/jphyseduc.v30i1.3054

Original Article

CONTENT VALIDITY, RELIABILITY AND CONSTRUCT VALIDITY OF A CHECKLIST FOR DIVE ROLL EVALUATION

VALIDADE DE CONTEÚDO, CONFIABILIDADE E VALIDADE DE CONSTRUTO DE UMA LISTA DE CHECAGEM PARA AVALIAÇÃO DO ROLAMENTO PEIXE

Rafaela Zortéa Fernandes Costa[1], **Josiane Medina-Papst**[1], **Rodrigo Martins de Oliveira Spinosa**[1], **Dalberto Luiz de Santo**[1] and **Inara Marques**[1]

[1]State University of Londrina, Londrina-PR, Brazil.

RESUMO

O objetivo deste estudo foi verificar a validade de conteúdo, a confiabilidade e a validade de construto de uma lista de checagem para avaliação do Rolamento Peixe. Em relação à validade de conteúdo, 4 especialistas na área da Ginástica Artística avaliaram a lista de checagem utilizando-se da técnica Delphi. Os resultados demonstraram bons índices de validade de conteúdo para as questões avaliadas (IVC>0,80). Sobre a confiabilidade, seis avaliadores, analisaram 20 vídeos do rolamento peixe de 10 crianças, utilizando-se da lista de checagem. Os resultados demonstraram concordância intra (ICC>0,80) e inter avaliadores (ICC>0,90). Por fim, para a validade de construto, 156 vídeos da habilidade foram analisados utilizando a lista de checagem. Foram aplicados os testes Alfa de *Cronbach* e Análise Fatorial Exploratória com rotação *Varimax*, a qual indicou boa consistência interna (α=0,76) e a existência de três fatores que explicam a lista de checagem. Desta forma, conclui-se que a lista de checagem é válida e confiável para a avaliação da habilidade motora proposta.

Palavras-chave: Aprendizagem. Destreza motora. Reprodutibilidade de testes. Checklist.

ABSTRACT

The objective of this study was to verify the content validity, reliability and construct validity of a checklist for dive roll evaluation. As for content validity, 4 experts in the Artistic Gymnastics field evaluated the checklist using the Delphi technique. Results showed good content validity indexes for the evaluation questions (CVI>0.80). About reliability, 6 raters analyzed 20 dive roll videos of 10 children, using the checklist. Results showed intra-rater ((ICC>0.80) and inter-rater agreement (ICC>0.90). Finally, concerning construct validity, 156 skill videos were analyzed using the checklist. Cronbach's alpha and Exploratory Factor Analysis with Varimax rotation were applied, with the latter indicating good internal consistency (α=0.76) and the existence of three factors that explain the checklist. Thus, it is concluded that the checklist is valid and reliable to evaluate the proposed motor skill.

Keywords: Learning. Motor Skills. Test Reproducibility. Checklist.

Introduction

Planning methods for any research requires the adoption of instruments and the development of procedures that ensure reliable results for the investigated phenomenon. Thus, researchers from different areas of knowledge must be attentive to the selection of instruments that have validity and reliability. In this line of thought, the construction and validation of checklists for qualitative evaluation of motor skills may help researchers in the Physical Education field, since it would ensure a more objective evaluation on the execution of the motor skill to be learned. Moreover, a valid and reliable instrument may help in many other contexts, such as those linked to practice, providing a support to physical education teachers and sport instructors.

It is understood that scientific development, as well as professional activity, especially in the Motor Learning field, lack measurement instruments that are valid and reliable. Thus, this research aimed to check the content validity, reliability and construct validity of a checklist for dive roll evaluation in artistic gymnastics (AG).

Said skill was chosen due to the amount and diversity of elements that it has, and to its distinction within the motor repertoire of most children. Broadly speaking, to execute rolling

skills, the head and the trunk must perform flexions and extensions in a coordinated manner to reach an angular movement[1]. Mastering the execution of this skill brings positive influences to a child's motor repertoire[2]. However, gymnastics skills are rarely practiced in school environments, especially because of their intrinsic difficulties, such as fear and lack of preparation, in addition to the scarcity of materials that teachers deal with to teach these skills[3], combined with the absence of a reliable instrument that may make it easier to track the acquisition of these skills during the learning process.

An instrument that evaluates the dive roll is the checklist designed by Medina-Papst[4], originally built with five steps: a) impulse; b) flight; c) landing; d) roll; and e) finish. Despite the relevance and possibilities that the instrument provides in the practice environment, it still lacks validation for its effective use. In addition to the lack of validity of the checklist in question, it is worth pointing out that no other instrument for motor skill evaluation in AG was found in the Brazilian literature, hindering the conduction of studies on the skills of the modality, as well as the practice of professionals, especially in educational contexts.

Understanding the importance that gymnastics skills have to motor development, as its acquisition would facilitate the learning of more complex skills, broadening an individual's performance possibilities[3], it is believed that the validation proposed by the study will help in professional practice, as well as in the conduction of further studies.

Nevertheless, it is known that determining the validity and reliability of an instrument is a complex process, and both are dependent on its purpose, on the interpretation of results that the instrument proposes, and on its use. In general, the concept of validity refers to the degree to which an instrument measures what is intended to be measured, and this guarantee for researchers may be checked by means of different pieces of evidence, namely: content validity, criterion validity and construct validity[5-7].

Construct validity allows verifying the meaning of the test[8]. It can be measured through the analysis of behavioral representation, which is basically done by means of two tests: factor analysis test and internal consistency analysis of the test[6-8]. To Pasquali[6], internal consistency analysis aims to verify the correlation of a test item with the other items or with the test overall result; the two tests most commonly used for said measurement are KR-20 and Cronbach's alpha[9]. In its turn, exploratory factor analysis seeks to verify how many common constructs are present in the instrument and to "evaluate" certain aspects of the phenomenon that is supposed to be studied, expressed by means of covariances[6].

Criterion validity refers to the instrument degree of efficacy in predicting an individual's performance. To measure this type of validity, the individual's performance in the instrument must correlate with the performance of the same individual when evaluated by another instrument that measures the same behavior (criterion)[6,8]. The great difficulty about checking this type of validity is the existence of valid tests that measure the same variable as the instrument that is supposed to be validated[6,7].

Finally, content validity seeks to verify the degree to which the instrument evidences a specific domain of content[5]. This validation step takes place in two phases: the phase for studying and constructing the instrument, and the phase when the instrument is evaluated by specialists[10,11].

Additionally to validity, it is important to verify the reliability of the instrument, which represents the coherence that the instrument has, shown by the consistency of results[5] over time (trustworthiness), or by means of the same rater (objectivity). A good indicator for the analysis of these assumptions (trustworthiness and objectivity) can be obtained through the intraclass correlation coefficient[12].

Methods

Because the process of validating instruments is a complex one, this study was conducted in three parts, namely: content validity; b) reliability; and c) construct validity. The research project was approved by the Ethics Committee on Research Involving Humans (Legal Opinion No 1.681.499 – CAAE 56871816.6.0000.5231).

Participants

Participants included 4 raters with expertise in AG and experience of at least 10 years in the modality. All of them were PhD professors and agreed to participate in the study voluntarily.

Reliability

To determine inter- and intra-rater correlation, 6 raters participated: three PhDs, one doctoral student, and two Masters of Physical Education. All of them agreed to participate in the study voluntarily.

As skill models, 10 children enrolled in the municipal education network of the city of Londrina, aged between 9 and 11 years old, participated as well. Four out of the 10 children had prior experience in AG, and six had no experience with the skill, which ensured the presence of both beginners and experienced individuals concerning the modality. Legal guardians and participants signed a free and informed consent form agreeing to participate in the study. The sample was composed by convenience, meeting three exclusion criteria, namely: having any diagnosed neurological impairment, taking medication for attention deficit and/or hyperactivity, or having any physical disability that would not allow the task to be performed.

Construct Validity

A total of 52 individuals – 30 children (10.5 years old ± 0.8) and 22 adults (23.6 years old, ± 2.3) – joined as skill models. All participants and their legal guardians agreed to participate in the study voluntarily and signed a free and informed consent form.

Instruments and Content Validity

The Google Forms platform was used for sending and receiving the questionnaire.

Reliability

To capture the images of the children that participated as models and executed the skill, large AG mattresses were placed on the schoolyard where collection happened. For filming, a digital Sony camera was used – Handycam DCR-SR42 model, sampling rate of 60 Hz, programmed with automatic shutter speed –, being positioned on the right side of the executor and in a perpendicular way to the execution of the roll.

Construct Validity

To execute the skill, large AG mattresses were placed on the schoolyard where collection happened. To capture the images, a digital Sony camera was used – Handycam DCR-SR42 model, sampling rate of 60 Hz, programmed with automatic shutter speed –, being positioned on the right side of the executor and in a perpendicular way to the execution of the roll.

Procedures

Content Validity

The instrument content was subjected to assessment by the raters, who answered a questionnaire using the Delphi technique. Seeking to organize the analysis of the answers, four evaluation components were defined for content validation (Figure 1).

Clarity of content description	Check if the language used for description and if the figure representing the phase or failures are of easy assimilation and/or comprehension by people who might use the instrument, teachers/professors and/researchers.
Technical pertinence	Check if the proposed phase or failures are adequate and pertinent for evaluating performance on the skill, and if the failure is in conformity with its corresponding phase.
Applicability	Check if the phase of movement failure can be observed by people who might use the checklist, and if it is important for teaching the skill.
Deduction	Check if the deduction values are representative of the failure presented.

Figure 1. Definitions of content validation evaluation
Source: The author himself

Thus, three questions were asked for each phase, and four questions were asked for phase failures, totaling 35 questions, with choices of answers on a Likert scale ranging from very good, good, regular, bad and very bad.

For each one of the five phases of the motor skill, there were three questions referring to (a) clarity of content description, (b) technical pertinence of the phase and (c) applicability of the list, totaling 15 questions. A second block with 20 questions explored failures in each phase of the skill, which concerned (a) clarity of content description, (b) technical pertinence of the phase, (c) applicability of the list and score for failure deduction. The last field of the instrument had a space for the raters to write general comments on the checklist.

Reliability

In front of a laptop, the participating children were instructed to watch twice the video demonstration of a skilled gymnast doing a dive roll and then execute the same skill three times on the mattress.

To measure the consistency and reproducibility of the criteria, the six raters analyzed the videos using the checklist modified by the AG experts during the content validity study. Each rater analyzed the same videos twice, with an interval of two weeks between observations. The videos were sent to the raters in slow motion via Google Drive. Thus, all of them watched the video at the same speed and as many times as they judged necessary. The checklist scores obtained during the evaluations were used for inter- and intra-rater correlations.

Construct Validity

In front of a laptop, each one of the 52 participants were instructed to watch twice the video demonstration of a skilled gymnast doing a dive roll and then attempt to execute the skill on a mattress three times. Afterwards, 3 raters (2 PhDs and 1 doctoral student in Physical Education) analyzed each individual's three attempts using the checklist, totaling 156 videos analyzed.

Data Analysis

Content Validity

Data analysis was carried out by means of Content Validity Index (CVI), defined by the sum of relative frequencies of answers 4 and 5 (good or very good), representing the raters' level of agreement on the evaluated items. A CVI higher than or equal to 0.80 was adopted – that is, most raters should judge the criterion as good or very good for it to be part of the final document –, as indicative of the adequacy of each instrument item[13].

Reliability

For data analysis, the Intraclass Correlation Coefficient (ICC) was applied to the scores of the raters' observations for each criterion on the checklist. The minimum agreement level adopted was 0.80[14].

Construct Validity

First, the Intraclass Correlation Coefficient (ICC) test was applied by means of the scores obtained in the observations to determine the raters' agreement, finding a good level of agreement (ICC=0.882; p=0.001). Thus, to verify the internal consistency of the test, Cronbach's alpha was used, presenting the total α of the instrument and, then, the α considering the exclusion of each instrument criterion, with $\alpha>0.70$[14] being adopted as minimum values. Later, an exploratory factor analysis was run by means of the analysis of main components with orthogonal rotation of Varimax type. The exploratory factor analysis comprehended: a) Kaiser-Meyer-Olkin (KMO) sample adequacy test with values higher than 0.50, and Bartlett's sphericity test, with $p<0.05$[14]; b) only factors with eigenvalues >1, according to Kaiser's rule[15]; c) factorial loads above 0.30 in the factor[6].

Results

For content validation, there were three evaluation rounds, in which several corrections and changes suggested by the raters were accepted, both in visual and written description and in the deduction values of each skill phase. At the end of the third round, the CVI values obtained for each criterion were higher than 0.80. Thus, the evaluation rounds were concluded. Table 1 displays results for the CVIs obtained with the raters' judgement.

Table 1. CVI of skill phases and failures according to the raters' judgement

Analyzed criteria	CVI for skill phases	CVI for skill failures
Clarity of description	0.90	0.95
Technical pertinence	1.0	0.85
List applicability	1.0	0.85
Deduction values	-	0.90

Source: The authors

The ICC identified a high level of agreement between the six raters (ICC=0.99; p=0.000). Thus, the test was applied again to determine values referring to the objectivity (inter-rater correlation) and trustworthiness (intra-rater correlation) of the instrument. As illustrated in Table 2, all levels of agreement concerning the instrument evaluation criteria obtained high agreement coefficients.

Table 2. Inter-rater correlation for reliability verification

Criterion	ICC	Confidence interval	*p*
Criterion 1	0.970	0.946 – 0.986	0.000
Criterion 2	0.961	0.930 – 0.982	0.000
Criterion 3	0.935	0.884 – 0.970	0.000
Criterion 4	0.893	0.840 – 0.951	0.000
Criterion 5	0.962	0.934 – 0.981	0.000
Criterion 6	0.858	0.746 – 0.934	0.000
Criterion 7	0.935	0.884 – 0.970	0.000
Criterion 8	0.942	0.896 – 0.973	0.000
Criterion 9	0.970	0.947 – 0.986	0.000
Criterion 10	0.997	0.974 – 0.994	0.000
Total	0.986	0.974 – 0.994	0.000

Source: The authors

When it comes to trustworthiness, all values showed high intra-rater agreement. Results are displayed in Table 3.

Table 3. Intra-rater correlation for reliability verification

Rater	ICC	Confidence interval	p
Rater 1	0.992	0.990 – 0.994	0.000
Rater 2	1.0	-	0.000
Rater 3	0.941	0.923 – 0.954	0.000
Rater 4	0.994	0.992 – 0.995	0.000
Rater 5	0.923	0.901 – 0.941	0.000
Rater 6	0.959	0.947 – 0.968	0.000

Source: The authors

Considering the results found, the instrument presents stable values over time (intra-rater correlation), as well as similar answers when analyzed by different raters (inter-rater correlation).

As for construct validity, Cronbach's alpha reliability coefficient showed satisfactory values in relation to the checklist in its totality (α=0.76). Furthermore, good α values were found when the instrument criteria were excluded from the test. Table 4 presents α values, considering the exclusion of criteria.

Table 4. Cronbach's alpha reliability coefficient, considering the exclusion of criteria

Criteria	α when the criterion is excluded
Jumping without bending the knees for impulse	0.739
Flexing the trunk, making the flight phase impossible	0.709
Not extending the legs	0.711
Keeping the arms flexed	0.739
Supporting the hands or head on the ground before the impulse, hindering the flight and the landing	0.715
Not putting the chin to the chest	0.764
Not executing this phase with the body tucked, extending the trunk	0.765
Rolling to the side and/or raising the trunk	0.758
Not finishing in tucked position	0.748
Using the hands to stand up	0.762

Source: The authors

Considering that all α values were higher than 0.70, the Exploratory Factor Analysis was applied, initially verifying the sample adequacy assumptions by means of the KMO test and Bartelett's sphericity test. The KMO value for the sample was 0.794, judged as good[16]. Likewise, Bertelett's sphericity test resulted in p=0.001, showing that the sample was adequate to the application of the Exploratory Factor Analysis[14]. Thus, analyses on eigenvalues and covariances for each criterion were run.

The test identified three factors with eigenvalues above 1, when Kaiser's criteria were established[15]. Altogether, these three factors explain 73% of the test variance. Table 5 summarizes these results.

Table 5. Factor matrix of the exploratory factor analysis concerning the checklist for dive roll evaluation

		Factors		
Criteria	Description of criteria	1	2	3
1	Jumping without bending the knees for impulse	0.645	-	-
2	Flexing the trunk, making the flight phase impossible	0.930	-	-
3	Not extending the legs	0.922	-	-
4	Keeping the arms flexed	0.840	-	-
5	Supporting the hands or head on the ground before the impulse, hindering the flight and the landing	0.882	-	-
6	Not putting the chin to the chest	-	0.303	0.804
7	Not executing this phase with the body tucked, extending the trunk	0.359	0.730	-
8	Rolling to the side and/or raising the trunk	-	0.536	0.505
9	Not finishing in tucked position	-	0.817	-
10	Using the hands to stand up	-	0.635	0.494
	Eigenvalues	3.951	2.231	1.198
	% of explained variance	39.514	22.306	11.982

Source: The authors

As seen in the table, criteria 1, 2, 3, 4 and 5 showed factorial loads in factor 1. Criteria 7, 8, 9 and 10 showed high factorial loads in factor 2, and criteria 6 showed higher factorial load in factor 3. Factorial load values below 0.30 (minimum established by Pasquali[6]) are represented by a dash.

Discussion

Broadly speaking, the objective of this study was to investigate the validity of the checklist initially designed by Medina-Papst[4] for dive roll evaluation. Thus, content validation, reliability and construct validation were analyzed. As for content validation, the checklist proved to be valid according to the AG experts' evaluation, with CVI below the critical threshold of 0.80[13] for all rated items. It is worth recalling that several changes have been made to the instrument in order to meet the demands presented by the raters. At the end of that step, it is understood that the checklist content was clear, applicable and pertinent for evaluating the proposed skill, and all deduction values are coherent with the severity of mistakes usually made during the execution of the skill.

The reliability analysis, carried out through the verification of inter- and intra-rater correlation, showed that the instrument, when used by researchers, provides constant results both over time and when employed by the same rater (Tables 2 and 3). According to Souza, Alexandre and Guirardello[17], reliability is one of the main criteria that shows the quality of an instrument, checking the extent to which the latter can reproduce consistent results over time

and by different observers, as done in this study. The investigation by Medina-Papst[4], which used the original instrument to evaluate the performance of 24 children on the skill, also showed correlation between the evaluation of three AG specialists, corroborating with the results of this study. These results indicate that the instrument is reliable and can be used by teachers/professors and/or researchers interested in employing the checklist in both educational and research contexts. About its use, it is worth stressing that, for being an instrument of easy application and low-cost, it is of great importance to the education context, in which teachers, even with little experience as to the skill or lack of equipment, can evaluate their students' skill level.

About construct validation, the test presented good internal consistency (α=0.76). Moreover, when considering the exclusion of criteria, all presented an α higher than 0.70, indicating the accuracy and association of criteria[7,16]

The exploratory factor analysis revealed the existence of three factors with eigenvalues above 1, as established by Kaiser's rule[15]. Altogether, they explain 73.80% of the test variance, higher than the minimum established by Pontes Júnior et al.[16] – 50%. Factor 1 showed an explanation power of 39.51% of the variance. Observing the checklist, it is possible to notice that the criteria that compose this factor refer to the Impulse, Flight and Landing phases. The phases discriminated in this factor, in addition to sequential, are pre-requisites to each other. For instance, there is no Flight phase if the executor does not perform the Impulsion. This indicates that the discrimination of these phases within factor 1 are representative of a common dimension. The explanation power of this factor (39.51%) reinforces the importance of these items to the test, which, in general, are composed of criteria that distinguish the dive roll skill from the forward roll basic skill. Additionally, criteria 7 ("not executing this phase with the body tucked, extending the trunk") showed a factorial load above 0.30 in this factor. It is understood that the factorial load observed in criteria 7, within factor 1, is due to the fact that the Flight Phase is performed with the body extended and that, possibly, the individual will have some trouble tucking his/her body in the next phase, the rolling.

Factor 2, composed of criteria concerning the Roll and Finish phases, presented an explanation power of 22.30% of the test variance. In general, these phases represent the roll basic skill, being of great importance to dive roll, since the roll basic skill is essential for the learning of this more complex motor skill[18]. Thus, when dive roll is performed, these phases are configured by the outcome of all previous phases. Therefore, these criteria have representativeness in factor 2. Criteria 6 ("Not putting the chin to the chest"), despite not being representative of factor 2, presented a substantial factorial load as well (0.303). Observing this criterion, it is possible to identify the great importance that it has to compose this factor, since it is a pre-requisite for success in the Roll and Finish phases.

Finally, factor 3 was composed of criterion 6 only. This criterion consists of observing if the individual is "Not putting the chin to the chest". Results show an explanation power of 11.98% of the test variance, which reveals the great importance of this criterion to compose the test. The high factorial load of this criterion, within an isolated factor, draws attention to the fact that it is a critical point during the execution of this skill. First, because it can be a good indicator for mature development of the roll basic skill, which is essential to dive rolling. Besides, not putting the chin to the chest is considered a typical error during dive rolling among little experienced or poorly-instructed executors[19,20]; in dive rolling, the correct position of the head during landing is an action that guarantees the individual's safety when executing the skill[12]. For this reason, it is worth stressing the importance of observing this criterion when evaluating the skill, as well as the proper guidance on this point in the execution of the skill.

Still regarding factor 3, it is possible to see that criteria 8 and 10 presented factorial loads in this factor. The covariance between these criteria and criteria 6 is representative of the factor because, if the individual performs the movement with his/her head positioned incorrectly, he/she will likely not be able to raise it properly, which will lead to failures in criteria 8 and 10.

Conclusions

As it has been made evident, the results of the present study confirm the content validity, reliability and construct validity of the instrument. These aspects are fundamental to the construction and application of evaluation instruments[6]. It is worth remembering that the literature in Brazil lacks instruments with validity and reliability[21], especially when it comes to specialized motor skills and ecological contexts.

Another point to be highlighted is the skill addressed in the study – dive roll –, which differs from most skills in the basic motor repertoire of people because, in addition to a full 180^{o} spin over the head, it requires the combination of a jump and a phase without contact with the ground (flight phase). From this perspective, it is understood that gymnastics skills are fundamental to motor development, as they facilitate the learning of more complex skills and broaden performance possibilities for different motor skills[3]. Thus, it is expected that the instrument may help researchers interested in investigating the learning of this skill, as well as teachers/instructors that need to evaluate their students' performance.

In conclusion, the proposed checklist is a valid and reliable instrument, presenting results within expected values. In this way, researchers and/or teachers/professors can use it in their evaluations. However, it is worth pointing out that, for being an instrument that evaluates a specific skill rarely approached in the literature, it was not possible to carry out criterion validity, since no valid instruments that rated the same skill were found. Besides, because this is an initial research, without previous empirical analysis of a database different from the current one, the Exploratory Factor Analysis was chosen to understand and explore the instrument constructs. However, a Confirmatory Factor Analysis should be run in the future, involving data sourced from new collections. Finally, further studies should be conducted, mainly with other populations, so as to confirm aspects related to the validity and reliability of this instrument.

References

1. Perotti Junior A, Pellegrini AM. Organização espaço temporal do rolamento para frente. Movimento 2001;7(15): 9-20. DOI: 10.22456/1982-8918.2619
2. Ramos ESH, Viana HB. A importância da ginastica geral na escola e sua importância para crianças e adolescentes. Movimento e percepção 2008,9(13):190-199.
3. Schiavon L, Nista-Piccolo VL. A ginástica vai a escola. Movimento 2007;13(3):131-150. DOI: 10.22456/1982-8918.3572
4. Medina-Papst J. Dicas de aprendizagem na aquisição do rolamento peixe por crianças com Transtorno do Desenvolvimento da Coordenação. [Dissertação de Mestrado em Educação Física]. Curitiba: Universidade Federal do Paraná; 2007.
5. Martins GA. Sobre confiabilidade e validade. Rev Bras Gest N 2006;8(20):1-12.
6. Pasquali L. Psicometria*:* teoria dos testes na psicologia e na educação. 3 ed. Petrópolis, RJ: Vozes, 2009a.
7. Pasquali L. Psicometria. Rev Esc Enferm USP 2009b;43(esp):992-999. DOI: 10.1590/S0080-62342009000500002
8. Raymundo VP. Construção e validação de instrumentos: um desafio para psicolinguística. Letras hoje 2009;44(3):86-93.
9. Urbina S. Fundamentos da testagem psicológica. Porto Alegre: Artmed; 2007.

10. Alexandre NMEC, Coluci MZO. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medida. Ciên Saúde Colet 2011;16(7):2061-2068. DOI: 10.1590/S1413-81232011000800006
11. Coluci MZO, Alexandre NMEC, Milani D. Construção de instrumentos de medida na área da saúde. Ciên Saúde Colet 2015; 20(3): 925-936. DOI: 10.1590/1413-81232015203.04332013
12. Thomas JR, Nelson JK, Silverman SJ. Métodos de pesquisa em atividade física. 6. ed. Porto Alegre: Artmed; 2012.
13. Davis LL. Instrument review: Getting the most from a panel of experts. Appl Nurs Res 1992;5(4):194-197. DOI: 10.1016/S0897-1897(05)80008-4
14. Field A. Descobrindo a estatística usando SPSS. 2. ed. Porto Alegre: Artmed; 2009.
15. Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas 1960;20(1):141-151.
16. Pontes Junior JAF, Soares ES, Almeida LS, Trompieri Filho N. Análise fatorial exploratória e alfa de cronbach: elementos iniciais na validação de instrumentos de avaliação educacional. Educação e linguagem 2014;1(1): 63-75.
17. Souza AC, Alexandre NMC, Guirardello EB. Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. Epidemiol Serv Saúde 2017;26(3):649-659. DOI: 10.5123/S1679-49742017000300022
18. Bortoleto MAC. Introdução a pedagogia das atividades circenses. Jundiai: Fontoura; 2008.
19. Santos JCE. Manual da ginastica olímpica. 2. ed. Rio de Janeiro: Sprint; 1985.
20. Leguet J. As ações motoras em ginástica esportiva. Editora Manole; 1987.
21. Nunes MÊS, Gehring PR, Basso L, Fonseca COM, Thomasi MG, Santos S. Construção e validação de um instrumento de análise qualitativa do arremesso (lance-livre) do basquetebol. Motriz 2012;18(4):627-635. DOI: 10.1590/S1980-65742012000400001

Acknowledgments: CNPQ and Araucária Foundation

Authors' ORCID:

Rafaela Zorite Fernandes Costa: 0000-0002-7634-7604
Josiane Medina-Papst: 0000-0002-9335-7898
Rodrigo Martins de Oliveira Spinosa: 0000-0002-3804-3239
Dalberto Luiz de Santo: 0000-0002-0860-0612
Inara Marques: 0000-0002-9554-9164

Received on Oct, 01, 2018.
Reviewed on Dec, 30, 2018.
Accepted on Feb, 20, 2019.

Author address: Rafaela Zortéa Fernandes Costa. Avenida Celso Garcia Cid., Pr 445, Km 380, Campus Universitário, Londrina PR, CEP [Postal code] 86057-970. E-mail: rafaela_zortea@hotmail.com.