International Journal of Aquatic Research and Education

Volume 13 | Number 4 Article 11

3-10-2022

Content, Construct, and Criterion Validity, Reliability, and Objectivity for Aquatic Readiness Assessment for Brazilian Children

Nadia C. Valentini
Universidade Federal do Rio Grande do Sul, nadiacv@esef.ufrgs.br

Keila R G Pereira
Universidade Federal do Rio Grande do Sul, keilargpereira@gmail.com

Glauber C. Nobre
- Instituto Federal de Educação, Ciência e Tecnologia do Ceará, glauber_nobre@hotmail.com

Recommended Citation
Valentini, Nadia C.; Pereira, Keila R G; and Nobre, Glauber C. (2022) "Content, Construct, and Criterion Validity, Reliability, and Objectivity for Aquatic Readiness Assessment for Brazilian Children," *International Journal of Aquatic Research and Education*: Vol. 13: No. 4, Article 11.
DOI: https://doi.org/10.25035/ijare.13.04.11
Available at: https://scholarworks.bgsu.edu/ijare/vol13/iss4/11

Abstract

The Aquatic Readiness Assessment (ARA) is an assessment instrument for measuring children's aquatic readiness. The objective of the study was to translate the English version into Portuguese and to investigate the content, construct, and criterion validity as well as the reliability and rater objectivity of the ARA for Brazilian children. Twenty-three professionals and 464 children, newborn to 13 years-old participated in the study. We found strong content (94% to 100% of judges' agreement) and criterion validity, internal consistency (α from .96 to .97), and inter-rater objectivity (ICC from .81 to .98), and test-retest reliability (ICC from .94 to .98). Appropriate fit indices were observed for the model (CFI = .99; TLI = .99; RMSEA .08, CI 90% = .67 to .10); the model was invariant for boys and girls (CFI = .99; RMSEA = .080; ΔCFI = .009; Δ RMSEA = .015) but not for age groups (CFI = .87, RMSEA = .160). The ARA presented adequate validity and reliability for evaluating the swimming performance of Brazilian children.

Keywords: validity, reliability, rater objectivity, water competence, aquatic readiness, aquatic developmental sequence patterns

Introduction

The acquisition of water competence has been recognized as essential to prevent drowning, especially among children. The WHO (World Health Organization) has highlighted that around the world, the highest drowning rates are among young children (1- to 4-years-old) followed by school-age children (5 to 9 years old). Moreover, in the western Pacific region, children aged 5- to 14-years-old die more frequently from drowning than any other cause; Globally drowning is one of the top five causes of death for individuals under 14-years-old in 48 of 85 countries that provided data meeting the WHO inclusion criteria (WHO, 2014). Therefore, acquisition of water competence, a crucial survival set of skills that reduce individual risks of drowning, is necessary.

Across different countries, the acquisition of basic water skills (e.g., back floating, breath control), strokes (e.g., front crawl, breaststroke), and safety procedures to prevent drowning and to create a foundation for learning more complex swimming skills (American Red Cross, 2009; McCool et al., 2008; Petrass et al., 2012; Stallman et al., 2008) has been the goal of aquatic programs. The importance placed on water competence is crucial to learn water safety and prevent drowning (Quan et al. 2015), but it is also a relevant content for child development (Courage, 2006; Erbaugh, 1986; Martins et al., 2010; Pan, 2010).

Water competence is a complex construct mediated by the constraints among swimmer's individual qualities, the goals and demands of each aquatic task, and the conditions associated with general and specific aquatic settings

(Langendorfer, 2015). Within this paradigm, aquatic readiness includes a unique set of aquatic fundamental skills and attitudes that precede the acquisition of more advanced aquatic skills (Langendorfer & Bruya, 1995). It may predict who is the most likely swimmer ready to learn more complex skills (Langendorfer, 2015; Stallman, et al., 2017). To implement appropriate programs to develop children's aquatic readiness and water competence, the right assessments are necessary.

Identifying children's water competence requires the use of reliable and validated tests with appropriate psychometric properties. An assessment has a fundamental role in understanding, establishing, and promoting motor development (Tamplain et al., 2020) and implementing appropriate intervention programs to improve motor competence (Burton & Miller, 1998). Assessment has been a challenge for many aquatic programs, precisely due to the different assessment goals (e.g., strokes or crucial fundamental skills) and the lack of validity evidence. For children, the literature provided examples of several assessments (i.e., Aquatic Skills Checklist; Erbaugh Rating Scale; Inventory of Evolutionary Aquatic Development; Humphries Assessment of Aquatic Readiness; Aquatic Readiness Assessment); however, little psychometrics evidence has been provided (Alaniz et al., 2017; Erbaugh, 1978; Langendorfer & Bruya, 1995; Pan, 2011). Besides, few assessments directly intend to measure water competence (Canossa *et al*., 2007; Costa et al., 2012; Langendorfer & Bruya, 1995; Wizer, Franken & Castro, 2016).

The Aquatic Readiness Assessment (ARA) is used to assess necessary skills that precede the acquisition of more advanced aquatic skills and water safety for children, grounded in the water competence model, and support teachers to plan effectively and timely activities (Langendorfer & Bruya, 1995). The ARA contains a set of nine developmental sequences of aquatic readiness – the water competence components (i.e., water entry, breath control, buoyancy and body position, arm actions, leg action, and combined movements) with specific developmental levels for each component. Some psychometrics were provided for the ARA regarding content developmental validity, intra- and inter-rater objectivity, and test-retest reliability. However, the ARA still lacks further establishment of additional relevant psychometric properties regarding content, construct, and criterion validity, reliability, and objectivity that support these instruments' use to assess children and childhood.

In addition, it is unknown whether the ARA which was developed for American children is suitable for children with different cultural backgrounds, such as Brazilians. Despite this restriction, the ARA has been used, for example, to examine intervention program effectiveness (Kjendlie & Mendritzki, 2012; Rocha et al., 2018) and to determine the optimal readiness for children’s advancement in an aquatic education program (Shannon, 2017). The contribution of these studies to the current knowledge of children's aquatic skills development is well recognized. Adequate assessment of children's motor competence depends on reliable and valid instruments using several psychometric approaches before using the test (Burton & Muller, 1998; Cronbach & Meehl, 1995; Vallerand, 1989; Yun & Ulrich, 2002). This study's objectives were to translate the Aquatic Readiness Assessment (ARA) from English to Portuguese and to investigate the content, construct, and criterion validity, reliability, and rater objectivity of the ARA among Brazilian children.

Method

Participants

A total of 23 professionals participated in the present study. Four bilingual translators, three professionals with Ph.D.s in human movement science and a focus in motor behavior and with extensive expertise in aquatics teaching, and 16 experienced professionals with majors in kinesiology and with aquatic teaching experience participated in the first phase of content validity in the present study. Two experienced aquatic teachers, doctoral candidates with majors in Kinesiology, were enrolled in the inter-rater objectivity process.

The sample of 464 children consisted of boys (n = 222) and girls (n = 242), ranging in ages from newborn to 13 years-old (newborn to 2-years-old: 69 children; 3- to 6-years-old: 150 children; 7- to 13-years-old: 245 children) from four cities located in different regions in Brazil. The sample was representative of Brazilian infants regarding socioeconomic status, gender, race, and age distribution. The children were recruited consecutively, with the permission of the institutions and parents. We obtained the consent from each child's parents, institutions, and from each professional who participated in the study. The university research ethics committee approved this study.

Instrument

Likert Scale for Language Clarity and Relevance for ARA Components

We developed a Likert scale to assess clarity and relevance of each ARA aquatic movement sequence with scores ranging from 1 to 5 (5 = highly clear/highly relevant; 4 = very clear/very relevant; 3 = more or less clear/more or less relevant; 2 = little clear/ little relevant; 1 = not clear/not relevant). We use the same scale to examine experts' and professionals' agreement (face validity).

Aquatic Readiness Assessment

The ARA is an individual observational assessment developed based on research and professional experience to assess basic components of children's aquatic readiness and water competence. The ARA contains developmental sequences of the basic aquatic movement patterns; a detailed description of each developmental sequence is also provided to illustrate a child's developmental level. For the water orientation and adjustment component, three developmental sequence levels were observed; for the water entry component, five ordered levels; for the breath control component, five progressive levels; for the buoyancy/flotation component, four developmental levels; for the body position component, four levels; for the arm propulsion action component, four levels; for the arm recovery action component, five levels; for the leg action component, five levels; and for the combined movement component, five levels. Table 1 summarizes the ARA components and developmental levels. A detailed description was presented of each developmental level for each component that a child could achieve as part of scoring each ARA component (Langendorfer & Bruya, 1995).

Evidence of validity and reliability has been provided for all the ARA components. Adequate indices were reported for developmental construct validity of each component (i.e., water orientation and adjustment, water entry, body position, arm propulsion action, arm recovery action, leg action, combined action); test-retest reliability (indices $\geq$ 90%; water orientation and adjustment, water entry, breath control, body position, arm propulsion action, arm recovery action, leg action, combined action); intra- and inter-rater objectivity (agreement $\geq$ 80%; water orientation and adjustment, water entry, breath control, body position, arm propulsion action, arm recovery action, leg action, combined action); and content validity for breath control and buoyancy components) (Cool, 1992, Roberton, 1977; Langendorfer, 1984a; Langendorfer et al., 1987, Balan & Langendorfer, 1988a, 1988b).

Procedures

Four bilingual translators independently enrolled in the double back-reverse independent translation (Hernandez-Nieto, 2002; Vallerand, 1989). Two independent translations of the ARA were conducted from English to Brazilian-Portuguese and two from Brazilian-Portuguese back to English. The translation included the component names and description, the developmental levels, the decision rules (i.e., the detailed description of each level), and the assessment guidelines. The content validity enrolled first the three experts who independently scored each ARA item's clarity and relevance using the 5-point Likert scale. Subsequently, 16 professionals received the ARA and used the same Likert scale to score all components' clarity and relevance.

Rater Objectivity

For the inter-rater objectivity, prior to beginning the study, raters A and B, experienced aquatic instructors, trained to use the ARA. The training consisted of studying the assessment, practicing the ARA with videos, conducting assessments, assessing children, recording assessments for posterior analysis, scoring children's

Table 1

Components and developmental levels of the Aquatic Readiness Assessment

Components	# Levels	Developmental Level Names
Water orientation and adjustment	3	1. No voluntary entry 2. Voluntary entry with hesitancy 3. Voluntary entry with no fear
Water entry	5	1. No voluntary entry 2. Assisted feet-first entry 3. Unassisted feet-first entry 4. Assisted head-first entry 5. Unassisted head-first entry
Breath control	5	1. Reflexive breath holding 2. Spitting or shipping 3. Voluntary face submersion 4. Repeated breath holding 5. Extended breath holding and/or rhythmic breathing with stroke
Buoyancy/flotation	4	1. No flotation 2. Flotation with assistance 3. Flotation with support 4. Unsupported flotation
Body position	4	1. Vertical 2. Inclined 3. Level 4. Horizontal
Arm propulsion action	4	1. No arm action 2. Short downward push 3. Long push-pull paddle 4. Lift propulsion
Arm recovery action	5	1. No arm action 2. No overwater recovery 3. Rudimentary overarm 4. Straight overarm 5. Bent-elbow overarm
Leg action	5	1. No leg action 2. Plantar push 3. Rudimentary flutter 4. Bent knee flutter 5. Straight-leg flutter
Combined movement	5	1. No locomotor behavior 2. Dog paddle 3. Beginner stroke 4. Rudimentary crawl 5. Advanced crawl

performance, and participating in meetings to discuss assessments. Each child's performances were scored by two raters individually in real-time for near 10% of the sample (N = 41 children). We conducted a test-retest reliability with a total sample of n = 464; children were assessed and re-assessed by the same professional within a one-week interval; we scored the children's performance in real-time.

Recruitment

We contacted the institutions, and one of the researchers explained the goals and procedures. Seven institutions, in four cities, agreed to participate and signed the institutional informed consent. We held a meeting with the teachers responsible for the water program to explain the assessment procedures. We contacted the parents and explained the research goal and procedures; parents who agreed with their child participating in the assessment signed the informed consent.

The assessments were conducted individually in a pool with comfortable water and air temperatures and in which the children were familiar. We conducted practice and formal trials for each item according to the children's responses and tolerance. During the testing, we elicited multiple trials (2 to 3) and under varied conditions to achieve the child's most advanced possible behavior. If the child was fearful, fatigued, or distressed, the testing was paused. For some developmental levels, the child was observed underwater using swim goggles or a mask, especially for younger children. The assessments were conducted in the corner of the pool to assess young children or children with little experience in the water. For infants and toddlers, parents entered the pool and assisted with the assessment procedures.

Data Analysis

We estimated the sample size using EpiInfo statistical software (version 7.0). Considering an approximate population of 200,000 children from four cities, 50% of expected frequency, 97% of confidence level, 4% marginal error, and 35% of possible attrition. A final estimated sample of approximately 660 children was necessary to achieve sufficient statistical power. We calculated descriptive analysis for all ARA sequences using the mean, standard deviation, skewness, kurtosis.

We conducted two different procedures to analyze the experts' and professionals' scores related to components' clarity and pertinence regarding content validity. The content validity coefficient (CVC) was first calculated with values > .70 considered as acceptable (Hernandez-Nieto, 2002). Second, Gwet's concordance coefficients test (AC1; Gwet, 2008) weighted by the scale's ordinal categories (Likert scale 1 to 5 for clarity and relevance) was conducted to complement the ICC analysis with values greater than .80 considered as adequately high agreement (Landis & Koch, 1977).

We examined possible multivariate outliers using Mahalanobis squared distance (D^2) and the Omnibus and Small's Chi^2 tests for the multivariate nonnormality of the data regarding construct validity. We conducted the Confirmatory Factorial Analysis (CFA) to examine ARA's relational structure, testing the models using weighted least squares mean and variance adjusted (WLSMV). We tested the overall fit of the model with the Tukey Lewis Index (TLI) and Comparative Fit Index (CFI) (Hair et al., 2010; Hu & Bentler, 1999); we accepted values greater than or equal to .95 and .90 as appropriate (Hair et al., 2010). We also used the Root Mean Square Error of Approximation (RMSEA) with a 90% confidence interval (CI 90%), adopting values lowest .05 and values between .06 and .08 as good and acceptable, respectively (Hair et al., 2010).

To verify the model invariant adjustment for sex and age groups (i.e., newborn- to 2-years-old, 3- to 6-years-old, and 7- to13-years -old) the invariance factorial analysis was loaded using Multigroup CFA. We conducted the configurational invariance analysis to determine if the number of components was the same for boys and girls and ages. We also used the metric invariance to verify if loadings varied across sex and age by groups and their relationships (Kline, 2011). We conducted the scalar invariance to analyze if the intercept terms for each variable and construct did not vary by groups. We compared the models using differences between constrained and unconstrained models, the delta of the RMSEA (Δ RMSEA), and CFI (Δ CFI), adopting the recommended cut-off ($< .015$) to support the invariance assumption. We assessed discriminant validity using the Heterotrait-Monotrait ratio (HTMT) of the correlations (Henseler, Ringlet & Sarstedt, 2015). Thresholds adopted were: .85 for strict and .90 for liberal discriminant validity (Henseler, Ringle & Sarstedt, 2014).

We calculated the alpha for ordinal data based on polychoric correlations to assess internal consistency. Values $\geq .70$ were considered acceptable (Farsen, Fiorini & Bardagi, 2017; Nunnally, 1978). Alternatively, the composite reliability (CR) was conducted (Fornell & Larcker, 1981). For this study, considering the number of components (9 components), the CR values equal or superior to .70 were considered adequate (Valentini & Damásio, 2016). The components' reliability also was assessed; values equal or superior to .25 were considered adequate (Hair et al., 2010).

We calculated one-way ANOVAs to examine the ARA item-developmental level validity with the Bonferroni *post hoc* test to verify the differences between groups if the age group were significant. We conducted all the analyses using AgreeStat2015.6 software, Mplus version 7.4 (Muthén & Muthén, 2012), and "Psych" package from R-free-software (Revelle, 2011); $p \leq .05$ was adopted.

We conducted an Intraclass Correlation Coefficient (ICC) to investigate test-retest reliability. The interpretation of the strength of ICC scores were adopted using recommended cut-offs (weak: ICC < .40; moderate: ICC between .40 and .59; strong: ICC between .60 and .74; very strong: ICC between .75 and 1.00; Cicchetti, 1994). A two-way mixed effect model, based on the mean of multiple measures, was used to examine the internal consistency (Qin, Nelson, McLeod, Eremenco & Coons, 2019; Shrout & Fleiss, 1979). In this study, the average variance extracted (AVE) was used as a measure of precision. Values superior to .50 were considered adequate (Valentini & Damásio, 2016). We examined the inter-rater objectivity using Intraclass Correlation Coefficients (ICC; Walters, 2009).

Results

Cultural Adaptation

After completing the four independent translations, all four professionals attended a meeting with two of the lead researchers for the study; in the meeting we compared all translated versions with the original English version of the ARA. The Brazilian-Portuguese versions were revised, and, upon unanimous agreement, a final translated and edited scale resulted in the Brazilian-Portuguese version of the ARA (ARA-BR).

Content validity

The results showed high concordance among the experts for the total components (CVC clarity from 98.4 to 100%; CVC relevance from 98.6 to 100%); the scores for all ARA components were very or totally clear and relevant. The AC1 coefficients of agreement (clarity: 0.94 to 1.00; relevance: 0.97 to 1.00) endorsed the experts' high agreement. The professionals' agreement was also high; ARA components were scored as very or totally clear (CVC values from 87 to 100% of agreement) and relevant (CVC values from 95 to 100% of agreement). Table 2 presents the CVC and AC1 for the clarity and relevance of the ARA components.

Construct Validity: Model Uni- and Bi-Dimensionality

Table 3 presents the mean, standard deviation, skewness, and kurtosis across groups. We found negative skewness for most of the components. The analysis for ARA's model structure examined unidimensional and bidimensional models (aquatic adjustment and locomotor dimensions).

Unidimensional Model

We excluded ten multivariate outliers detected in the D^2 test. The Omnibus test based on Small's test ($\chi2(18) = 2151.81$, $p < .001$) confirmed the multivariate nonnormality of data. The CFA results presented an adequate adjusts for CFI (.99) and TLI (.99), and acceptable adjustment for RMSEA (.08, CI 90% = .67 to .10).

We examine the invariance of the model for boys and girls and age groups using multigroup analysis. The model without constriction demonstrated configurational invariance for boys and girls (CFI = .99, RMSEA = .08). The loadings did not vary by sex (ΔCFI = .009; Δ RMSEA = .015). Nevertheless, the model indicated that the intercept terms for each variable and construct do not vary by sex (ΔCFI =.009; ΔRMSEA = .015). The model without constriction demonstrated configurational variance for age groups (CFI = .87, RMSEA = .160).

The analyses for the Newborn to 2-years-old and 3- to 7-years-old showed that the model configuration remained the same (CFI = .96, RMSEA = .08). For children 7- to 13-years-old the "water orientation and adjustment" component item did not remain on the model due to lack of variability (i.e., all children scored 3). The analyses also showed that the loadings varied for the Newborn to 2-years-old and 3- to 7-years-old models (ΔCFI = .04, ΔRMSEA = .015). For the 7- to 13-years-old children, the model showed adequate indexes (CFI = .98, RMSEA = .060). According to modified indices, we conducted correlations between measurement error in breath control and fluctuation, between breath control and body position, and fluctuation and body position components. Figure 1 presents the load factor for unidimensional models.

Bidimensional Model

A bidimensional model considering the aquatic adjustment and locomotor dimensions showed an adequate adjustment (CFI = .98, TLI = .98), but inadequate fit for RMSEA (.124). The combined movement in water item showed a lower and non-significant load factor (.01, p = .932). The modified indexes suggested a correlation between fluctuation and body position components. When the model was reanalyzed considering the modified indexes, the fit was acceptable (RMSEA = .07, CI 95%: = .04 to .09) to good (CFI = .99; TLI = .99). The multigroup analyses showed a configural (RMSEA = .08 CI 95%: = .04 to .08; CFI = .99), metric (ΔRMSEA = .01; ΔCFI = .01) and scalar invariance for sex (ΔRMSEA = .01; ΔCFI = .01). The model demonstrated a configural variance across age groups. The analyses also showed that the loadings of bidimensional models varied for newborn to 2-years-old and 3- to 6-years-old groups (ΔCFI = .88; ΔRMSEA = .18).

As with the unidimensional model, the water orientation and adjustment component was excluded from the 7- to 13-year-old model due to the lack of variability. The 7- to 13-year-old model showed an adequate adjustment for CFI and TLI (CFI = .98, TLI = .98), but inadequate fit for RMSEA (value = 1.06). The modified indexes suggested a correlation between fluctuation and arm position action and arm recovery action and combined movement sequence movement patterns. The model was reanalyzed, and the indexes became good to acceptable (RMSEA = .08; CFI = .99; TLI = .99). The discriminant validity analysis assessed through Heterotrait-Monotrait ratio (HTMT) of the correlations shown an

Table 2

Content Validity Coefficient (CVC) and Gwet's Agreement Coefficients (AC_1) for language clarity and relevance for ARA components.

Experts (n=3)	Components' Clarity			Components' Relevance		
	CVC (%)	AC_1(IC 95%)	*p*	CVC (%)	AC_1 (IC 95%)	*p*
E1 × E-2 × E-3	98.6	-	-	99.6	-	-
E1 × E-2	98.9	.97 (.92 to 1.0)	< .001	100	1.0 (1.0 to 1.0)	-
E1 × E-3	100	1.0(1.0 to 1.0)	-	100	1.0 (1.0 to 1.0)	-
E2 × E-3	98.4	.94 (.86 to 1.0)	< .001	98.6	.97 (.92 to 1.0)	< .001

Note. E1: Expert 1; E2: Expert 2; E3: Expert 3; IC: Interval of Confidence.

Table 3

Mean, Standard Deviation, Skewness, and Kurtosis for ARA components

Components	Mean (Standard Deviation) & Skewness and Kurtosis n = 464											
	Boys		Girls		NB- to 2-y-old		3- to 6-y-old		7- to 13-years-old		Total Sample	
	M(SD)	Sk/Kt	M(SD)	Sk/Kt	M(SD)	Sk/Kt	M(SD)	Sk/Kt	M(SD)	Sk/Kt	M(SD)	Sk/Kt
Water Ori & Adj	2.7(.66)	-2.3/3.2	2.9(.50)	-3.6(4.3)	1.8(1.0)	.30(-1.9)	3.0(.20)	-6.3(6.3)	3.0(0)	-	2.8(.60)	-2.8(.60)
Water entry	3.7(1.4)	-.40/-1.2	3.6(1.3)	-.20(-1.3)	1.9(1.0)	1.0(.70)	3.8(1.1)	-.10(-1.6)	4.0(1.2)	-.70(-1.2)	3.6(1.4)	-.40(-1.3)
Breath control	3.8(1.6)	-.90/-.80	3.9(1.5)	-1.0(-.70)	2.1(1.2)	.90(-.10)	4.4(1.0)	-1.8(3.1)	4.0(1.6)	-1.2(-.40)	3.8(1.6)	-.90(-.70)
Buoyancy/flotation	3.4(.90)	-1.3/.40	3.5(.90)	-1.2(.10)	2.6(.80)	-.40(-.10)	3.7(.70)	-1.8(1.8)	3.6(.90)	-1.6(.90)	3.5(.90)	-1.3(.20)
Body position	2.8(1.2)	-.30/-1.5	2.8(1.2)	-.40(-1.4)	1.3(.70)	2.7(3.0)	3.2(.90)	-.90(-.20)	3.0(1.2)	-.60(-1.2)	2.8(1.2)	-.40(-1.4)
Arms propulsion	2.3(1.0)	.10/-1.1	2.4(1.0)	0(-1.2)	1.2(.60)	3.4(3.1)	2.3(.70)	.10(-.30)	2.7(1.0)	-.40 (-1.0)	2.3(1.0)	.10(-1.1)
Arms recover	2.7(1.4)	.20/-1.2	2.7(1.4)	.20(-1.1)	1.3(.80)	3.9(3.2)	2.6(.90)	0(-.40)	3.3(1.4)	-.30(-1.1)	2.7(1.4)	.20(-1.5)
Legs actions	2.9(1.5)	.20/-1.4	3.1(1.5)	-.10(-1.5)	1.6(.80)	2.4(4.1)	2.9(1.2)	.10(-1.1)	3.5(1.5)	-.50(-1.3)	3.0(1.5)	0(-1.4)
Combined Mov.	2.8(1.5)	.10/-1.4	2.8(1.4)	.10(-1.3)	1.3(.70)	3.9(5.2)	2.7(1.0)	-.10(-.90)	3.3(1.4)	-.50(-1.1)	2.8(1.4)	.10(-1.4)

Note. NB: Newborn; y-old: years-old; Sk: Skewness; Kt: Koutosis; Water Ori. & Adj.: Water orientation and adjustment; Combined Mov: Combined movement.

Figure 1

ARA-BR unidimensional models (Model 1: Total children; Model 2: Newborn to 2-years-old; Model 3: 3- to 6-years-old; Model 4: 7- to 13-years-old).

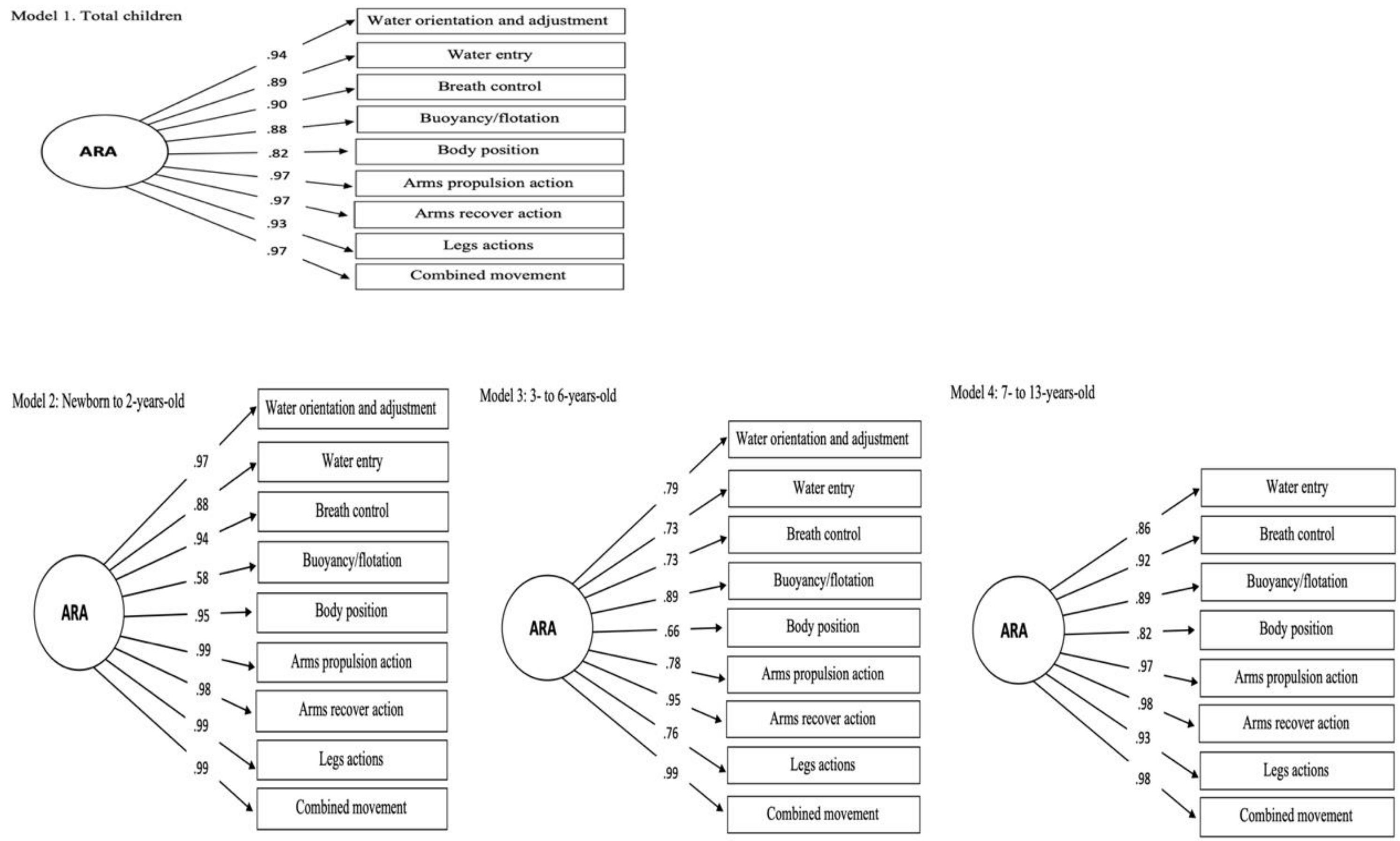
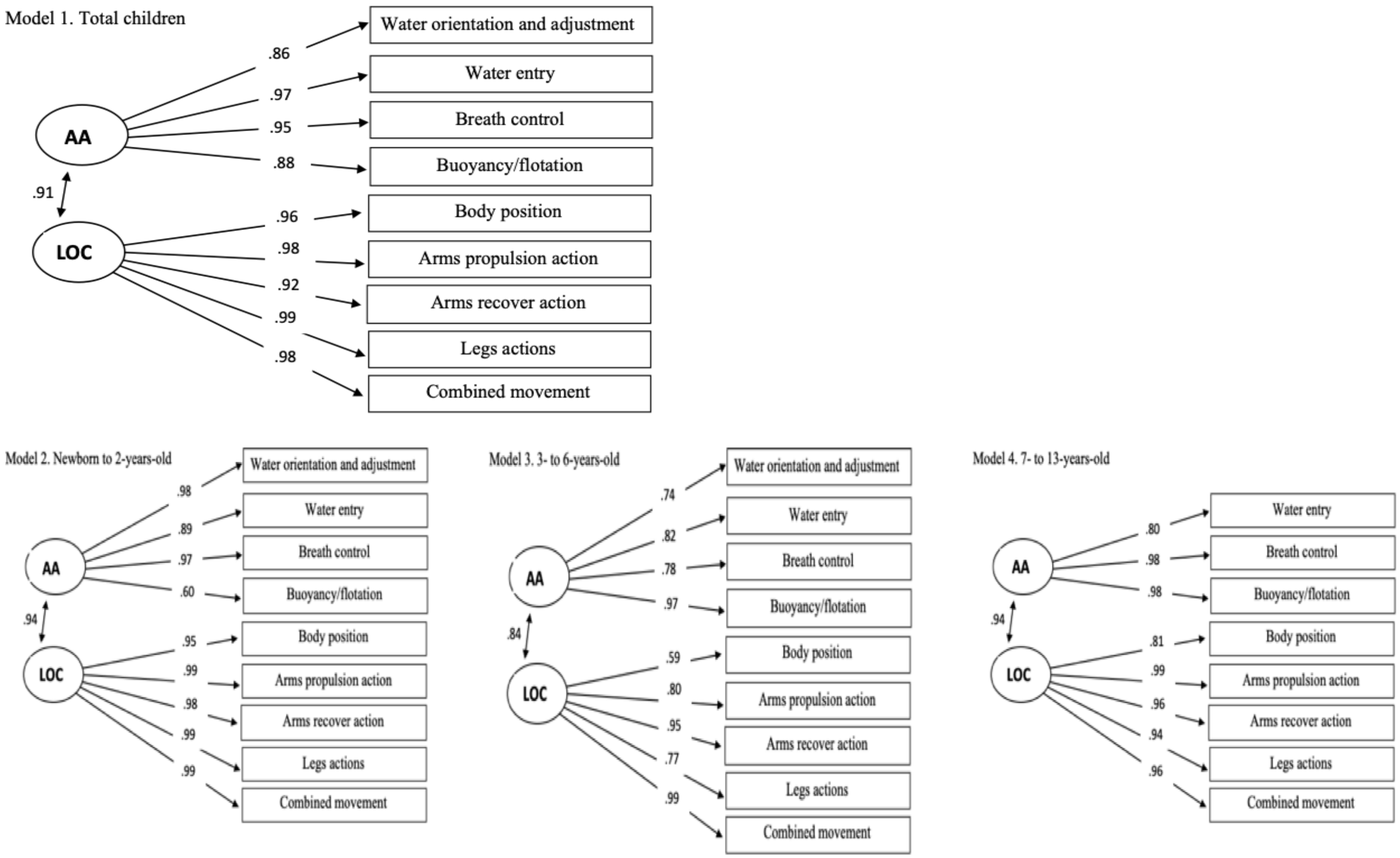


Figure 2

ARA-BR bidimensional models for aquatic adjustment (AA) and locomotor (LOC): Model 1: Total children; Model 2: newborn- to 2-years-old; Model 3: 3- to 6-years-old; Model 4: 7- to 13-years-old).



DOI: https://doi.org/10.25035/ijare.13.04.11

Table 4

ARA-BR internal consistence, test-retest reliability, and inter-rater objectivity

Components (n = 464)	Item-Test	α[1]	Test-retest Reliability ICC (95% CI)	Inter-rater Objectivity ICC (95% CI)
Water Ori. & Adj.	.80	.97	.96 (.95 to 97)	.97 (.90 to .99)
Water entry	.90	.96	.97 (.96 to .98)	.97 (.90 to .99)
Breath control	.90	.96	.96 (.95 to 97)	.87 (.60 to .96)
Buoyancy/flotation	.86	.97	.96 (.95 to .97)	.81 (.40 to .94)
Body position	.82	.97	.94 (.93 to .95)	.96 (.88 to .98)
Arms propulsion action	.93	.96	.97 (.96 to .98)	.98 (.93 to .99)
Arms recover action	.97	.96	.97 (.96 to .98)	.98 (.94 to .99)
Legs actions	.94	.96	.97 (.96 to .98)	.97 (.92 to .99)
Combined movement	.91	.96	.98 (.97 to .98)	.98 (.94 to .99)
Total	-	.97	-	-
Total Score (sum of components)	-	-	.99 (.98 to .99)	.98(.94 to .99)

Note. α: alpha coefficient; [1]If item dropped; Item-rest: polyserial correlation between the item and the sum of the rest of the item scores; Water Ori. & Adj.: Water orientation and adjustment inadequate discriminant validity between aquatic adjustment and locomotor dimensions in all tested models (Total children = .90; Newborn to 2-years-old = .98; 3- to 6-years-old = .91; 7- to 13-years-old = .98). Figure 2 presents the load factor for bidimensional models.

Construct Validity: Internal Consistency

The alpha coefficient from polychoric correlations showed appropriate values among components (α values from .96 to .97), for the total scale ($\alpha = .97$). Individual item's reliability (values from .67. to .94), the composite reliability (value = .98) and the average variance extracted AVE results (value = .85) also were appropriate. Table 4 presents the ARA-BR results for internal consistency, test-retest reliability, and inter-rater objectivity.

Developmental Criterion Validity

The one-way ANOVA showed a significant age group effect for all components (aquatic orientation and adjustment: $F(2,461) = 243.27$, $p < .001$; water entry: $F(2,461) = 86.61$, $p < .001$; breath control: $F(2,461) = 69.34$, $p < .001$; Buoyancy/flotation: $F(2,461) = 86.13$, $p < .001$; arm propulsion action: $F(2,461) = 73.63$, $p < .001$; arm recovery action: $F(2,461) = 76.47$, $p < .001$; legs actions: $F(2,461) = 53.94$, $p < .001$; combined movement: $F(2,461) = 76.51$, $p < .001$). Bonferroni *post hoc* tests showed that the Newborn to 2-years-old group scored significantly lower in all aquatic components compared to other groups (p values < .001). Children 3- to 6-years-old showed significant lower scores in breath control ($p = .047$), arm propulsion ($p < .001$), arm recovery ($p < .001$), legs actions ($p < .001$) and combined movement ($p < .001$) than 7-to 13-years-old.

Figure 3 presents the ARA-BR aquatic sequence patterns scores: Water orientation (3a), water entry (3b), breath control (3c), and buoyancy/flotation (3d) by age groups (** $p < .001$; * $p < .005$).

Figure 4 presents the ARA-BR aquatic sequence pattern scores: Body position (4a), arm propulsion action (4b), arm recovery action (4c), leg actions (4d), and combined movement (4e) by age groups (** $p < .001$; * $p < .005$).

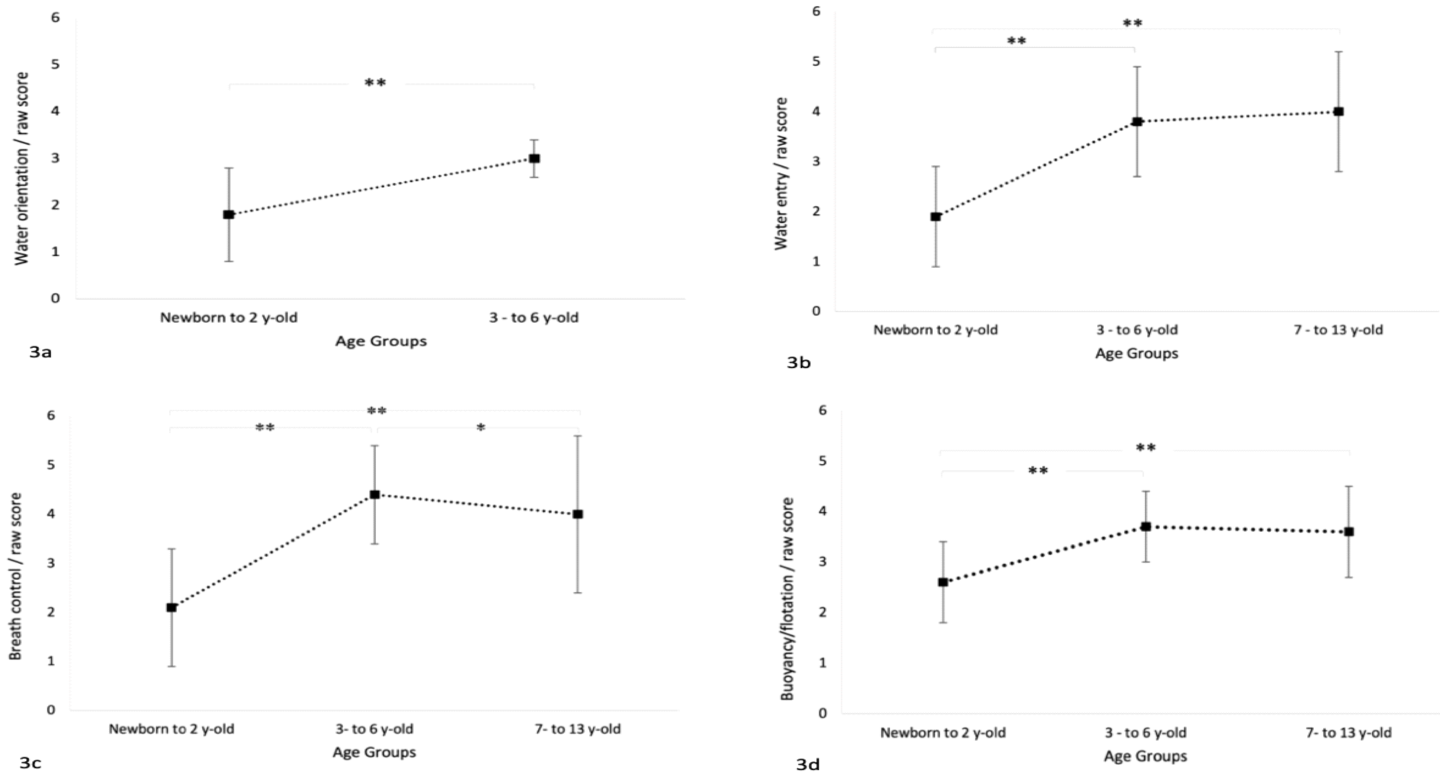
Inter-Rater Objectivity and Test-Retest Reliability

The ICC analysis showed high inter-rater objectivity (ICC from .81 to .98; *CI* 95% = .40 to .99). The test-retest reliability analysis showed high interclass coefficient correlation for ARA-BR components (ICC from .94 to .98; CI 95% = .93 to .98). Table 4 shows the α coefficient from ordinal data based on the components' polychoric correlations, item-test correlation, and ICC analyses (see Table 4).

DOI: https://doi.org/10.25035/ijare.13.04.11

Figure 3

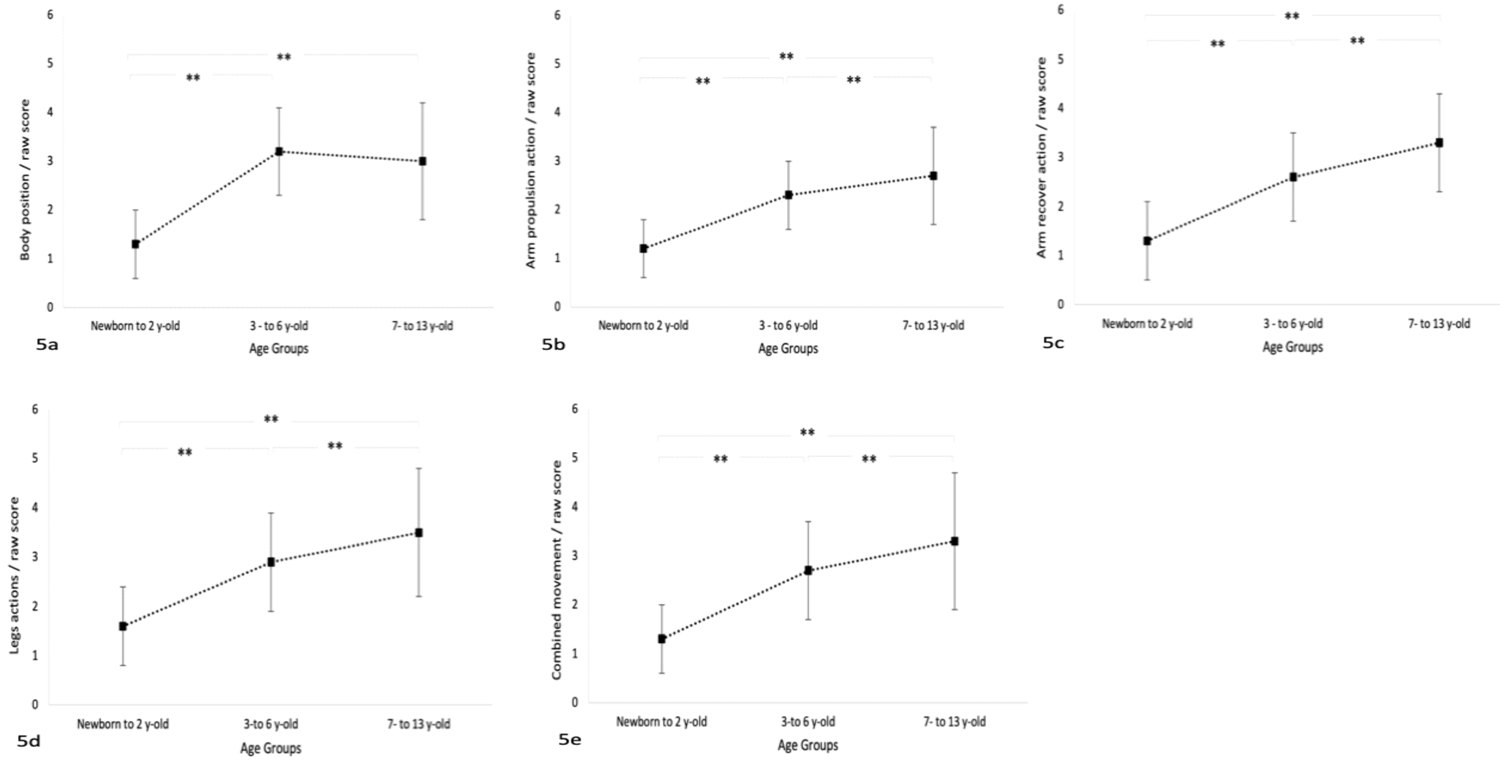
*ARA-BR: 3a water orientation (max. score 3), 3b water entry (max. score 4), 3c breath control (max. score 5), and 3d buoyancy/flotation (max. score 4) scores by age groups; ** p < .001; * p < .005.*



** $p < .001$; * $p < .005$

Figure 4

*ARA-BR: 4a body position (max. score 4), 4b arm propulsion (max. score 4), 4c arm recovery (max. score 5), 4d leg action (max. score 5), and 4e combined movement (max. score 5) scores by age; groups; ** p < .001; * p < .005.*



** p < .001; * p < .005

Discussion

The present study's objective was to translate and investigate the reliability and content, construct, and criterion validity of the Aquatic Readiness Assessment for Water Competence among Brazilian children.

ARA Translation and Cultural Adaptation

The independent translations resulted in a unified and final Brazilian-Portuguese version, the ARA-BR. All processes adopted in double-back and reverse translation and the committee meeting diminished the subjectivity influences that usually occur when an instrument is translated for another culture by only one translator (Vallerand, 1989; Hernandez-Nieto, 2002).

Content Validity: Experts and Professionals' Agreement

The content validity measures how test scores reproduce a specific construct (Furr, 2018). The CVC results and the AC1 tests indicated strong agreement (Cronbach & Meehl, 1976) among the three experts and the 16 professionals. The results showed that the ARA's components are comprehensible and relevant to assess the aquatic developmental sequence patterns and readiness in infants, toddlers, and children from 3- to 13-years-old. The results emphasized the proper representation of the components related to scale concepts and the test's theoretical relevance (Cronbach, 1989; Hernandez-Nieto, 2002). It is also important to emphasize that the content validity process enrolled experts and professionals with aquatic teaching experience, however, with different knowledge and skills. Therefore, the agreement regarding the ARA content by those audiences reinforces this assessment's potential by professionals with different degrees of experience.

Construct Validity: ARA-BR Models Structure

The inadequate RMSEA for the bidimensional model and the high correlations among components from aquatic adjustment and locomotor dimensions indicated that the unidimensional model better represents ARA's relational structure. The evidence from the discriminant validity analysis revealed that the two dimensions were nearly indistinguishable (Henseler, Ringlet & Sarstedt, 2015). Using the assessment as an essential part of the pedagogical practice (Langendorfer & Bruya, 1995) should provide aquatic practitioners with information regarding each component of the aquatic developmental sequence pattern; a child’s performance on each component could be different levels of the learning process (Wizer, Franken & Castro, 2016). The instructor can direct specific activities for each skill that is being practiced, understand the component with which the child is having difficulty, and provide specific instruction and feedback related to the difficulty.

Internal Consistency

The ARA components showed high internal consistency with a high Cronbach's alpha coefficient (above .96). Previous earlier studies also provided evidence for the validity of developmental sequence instruments such as the ARA (Cool, 1992; Roberton, 1977; Langendorfer, 1984a; Langendorfer et al., 1987; Balan & Langendorfer, 1988a, 1988b). Alpha levels above .60 are acceptable (Cronbach & Meehl, 1976); in the present study Cronbach alpha levels were all above .80. It is vital to notice that it is challenging to obtain this magnitude of coefficient in large samples (Cronbach & Meehl, 1976). Therefore, the present study results were outstanding, revealing a high homogeneity profile in the components and provided further evidence that the components are measuring the same construct (e.g., aquatic readiness or water competence) (Breakwell, Hammond, Fife-Schaw & Smith, 2006).

Developmental Criterion Validity

The ANOVA results confirmed the validity of the ordered developmental levels; we found significant changes across ages for each of the aquatic sequence components. Younger children showed lower developmental levels of aquatic readiness than the middle and older children, providing evidence for the ARA's component-developmental level criterion validity. In addition, we found that for the components of water entry (maximum score 4), buoyancy/flotation (maximum score 4), and body position (maximum score 4) high stability in the scores existed from among the older groups (3- to 6-years-old and 7- to 13-years-old). The original authors developed the ARA under the developmental assumptions that a developmental assessment must detect and measure significant age-related qualitative motor behavior changes and that within a reliable and valid developmental sequence for each component, most children show the same progression of change across levels. Therefore, we expected that the behaviors listed within ARA should regularly be observed in children, and the aquatic sequences should change in a robust order over time, and at any point of time they would be stable and consistent (Langendorfer & Bruya, 1995). We in fact observed this in the present study. Therefore, the ARA was able to differentiate among children with different motor responses across ages, showing itself to be an appropriate test for assessing the developmental level of aquatic skills across a broad age group from newborn to age 13 years.

Inter-Rater Objectivity and Test-Retest Reliability

We found high reliability which is the consistency in the score responses of a test (Furr, 2018) in the present study for test-retest procedures and for inter-rater objectivity. We also found strong agreement among raters (ICC values: total ARA above .98; item values from .81 to .98), indicating that the ARA-BR is reliable (Walters, 2009). The inter-rater results in the present study indicated a substantial agreement among evaluators (Landis & Koch, 1977; Vallerand, 1989) and were similar to those reported previously by the authors of the test for American children, (i.e., agreement ≥ 80% for water entry, body position, arms propulsion action, arms recovery action, leg actions, and combined stroke actions (Langendorfer & Bruya, 1995).

We conducted the test-retest reliability to investigate the temporal stability of the test (Cicchetti & Rourke, 2004), and we found high scores (above .94) for the assessments conducted within a one-week interval, suggesting a robust indication of test reliability (Waltz et al., 2010). Test-retest reliability was reported by the test authors with scores also above .90 for several aquatic sequence patterns (i.e., water entry, body position, arm propulsion actions, arm recovery actions, leg actions, combined stroke actions); Langendorfer & Bruya, 1995). Here, we advanced upon these earlier studies to have provided evidence for all nine of the aquatic sequences.

Temporal stability is related to the construct of a test and, therefore, it is an essential primary measure of a test's psychometric properties (Cicchetti & Rourke, 2004); the high levels attained in our study suggested that the ARA is reliable for assessing children over time, and therefore crucial to making teaching decisions regarding aquatic program activities. More recent aquatic pedagogical approaches have reduced the focus on teaching the swimming styles (i.e., strokes) (Lobo Costa, 2010) and to focus on the child’s interactions with the water to improve knowledge and survival skills. The ARA offers elements for working beyond swimming styles and a whole range of skills that allow children to feel safe and enjoy the water (Quan et al., 2015).

Conclusion

The results supported the content validity concerning clarity and pertinence by experts and professionals. The inter-rater scores were high and positive, confirming the ARA rater objectivity. We found positive and significant associations among the test and re-test assessments; the scale showed temporal stability for the total sample. The indices of fit for the unidimensional model were all appropriate and better represented the ARA construct. The internal validity results suggested high homogeneity for the ARA components, providing evidence that the ARA components embody the same construct. The ARA scores for developmental level criterion validity showed relevant evidence for the practical repercussions regarding different groups of children, and appropriate program strategies. Further longitudinal and concurrent evidence is still necessary to be demonstrated for the ARA, which is a limitation in the present study, and is included among our recommendation to future studies.

References

Alaniz, M. L., Rosenberg, S.S., Beard, N. R., & Rosario, E. R. (2017). The effectiveness of aquatic group therapy for improving water safety and social interactions in children with autism spectrum disorder: A pilot program. *Journal of Autism and Developmental Disorders, 47*(12), 4006-4017. https://doi.org/10.1007/s10803-017-3264-4X

Burton, A. W., & Miller, D. E. (1998). *Movement Skill Assessment*. Human Kinetics.

Breakwell, G. M., Hammond, S., Fife-Schaw, C., & Smith, J. A. (2006). *Research methods in psychology* (3rd ed.). Sage Publications.

Canossa, S., Fernandes, R.J., Carmo, C., Andrade, A., Soares, S., & Soares, S. M. (2007). Ensino multidisciplinar em natação: Reflexão metodológica e proposta de lista de verificação. *Motricidade, 3*(4), 82-99. https://Doi.org/10.6063/motricidade.656

Courage, M.L., Reynolds, G.D., & Richards, J. (2006). Infants attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development*, *77,* 680-695. https://Doi.org/10.1111/j.1467-8624.2006.00897.x

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological test. *Psychological Bulletin*, *52*(4), 281–302. https://Doi.org/10.1037/h0040957

Cronbach, L.J. (1989). Construct validity after thirty years. In R. L. Linn (Ed.), *Intelligence: Measurement, theory and public policy*. *Proceedings of a symposium in honor of Lloyd G. Humphreyes* (pp.147-171). University of Illinois Press.

Erbaugh, S. J. (1978). Assessment of swimming performance of preschool children. *Perceptual and Motor Skills, 47*(3), 1179-1182. https://Doi.org/10.2466/pms.1978.46.3f.1179

Erbaugh, S. J. (1986). Effects of aquatic training on swimming skill development of preschool children. *Perceptual and Motor Skills, 62*(2), 439-446. https://Doi.org/10.2466/pms.1986.62.2.439

Furr, R. Michael (2018). *Psychometrics: An introduction*. SAGE Publications.

Gwet, K. L. (2008). Variance estimation of nominal-scale inter-rater reliability with random selection of raters. *Psychometrika*, *73*(3), 407–430. https://Doi.org/10.1007/s11336-007-9054-8

Halverson, L. & Williams, K. (1985). Developmental sequences for hopping over distance: A prelongitudinal screening. *Research Quarterly for Exercise and Sport*, *56* (1), 37–44. https://Doi.org/10.1080/02701367.1985.10608429

Hernandez-Nieto, R. (2002). *Contributions to statistical analysis.* Los Andes University Press.

Kjendlie, P. L., & Mendritzki, M. (2012). Movement patterns in free water play after swimming lessons with flotation aids. *International Journal of Aquatic Research and Education, 6*(2), 149-155.

https://Doi.org/10.25035/ijare.06.02.06
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174. https://Doi.org/10.2307/2529310
Langendorfer, S. J. (2015). Changing learn-to-swim and drowning prevention using aquatic readiness and water competence. *International Journal of Aquatic Research and Education*, *9*(1). 1-8. https://Doi.org/10.25035/ijare.09.01.02
Langendorfer, S. & Bruya, L. (1995). *Aquatic readiness: Developing water competence in young children*. Human Kinetics.
Lobo da Costa, P. H. Pedagogia da natação: uma revisão sistemática preliminar. *Revista Mackenzie de Educação Física e Esporte*, *9*(1), 50-54. https://Doi.org/10.1590/S1807-55092006000100001
Martins, M., Silva, A.J., Marinho, D., Barbosa, T., & Sarmento, P. (2010). Assessment of heart rate during infants' swim session. *International Sports Medicine Journal*, *11*, 336- 344. https://Doi.org/10520/EJC48397
McCool, J.P., Moran, K., Ameratunga, S., & Robinson, E. (2008). New Zealand beachgoers swimming behaviors, swimming abilities, and perception of drowning risk. *International Journal of Aquatic Research and Education, 2*, 7–15. https://Doi.org/10.25035/ijare.02.01.02
Pan, C-Y. (2010). Effects of water exercise swimming program on aquatic skills and social behaviors in children with autism spectrum disorders. *Autism*, *14*(1), 9-28. https://Doi.org/10.1177/1362361309339496
Pan, C-Yu. (2011). The efficacy of an aquatic program on physical fitness and aquatic skills in children with and without autism spectrum disorders. *Research in Autism Spectrum Disorders*,*5*(1), 657-665. https://Doi.org/10.1016/j.rasd.2010.08.001
Petrass, L.A., Blitvich, J.D., McElroy, G.K., Harvey, J., & Moran, K. (2012). Can You Swim? Self-report and actual swimming competence among young adults in Ballarat. *Australia. International Journal of Aquatic Research and Education, 6*, 136–148. https://Doi.org/10.25035/ijare.06.02.05
Quan, L., Ramos, W., Harvey, C., Kublick, L., Langendorfer, S. J., Less, T. A., Fielding, R. R., Dlake, S., Barry, C., Shook, S., Wernicki, P. (2015). Toward defining water competency: An American Red Cross definition. *International Journal of Aquatic Research and Education, 9*, 12-23. https://Doi.org/10.1123/ijare.09.01.02
Roberton, A.A., & Halverson, L.E. (1988) The development of locomotor coordination. *Journal of Motor Behavior, 20*(3), 197-241. https://Doi.org/10.1080/00222895.1988.10735443
Rocha, H. A., Marinho, D. A.; Garrido, N. D., Morgado, L. S., & Costa, A. M (2018). The acquisition of aquatic skills in preschool children: Deep

versus shallow water swimming lessons. *Motricidade, 14*(1), 66-72. https://Doi.org/10.6063/motricidade.13724

Shannon, T. (2017). A Private vs. Community Aquatic Programming Comparison. *Oklahoma Association for Health, Physical Education, Recreation and Dance, 54*(2), 74-88. https://ojs.library.okstate.edu/osu/index.php/OAHPERD/article/view/7233/6662

Stallman, R., Junge, M., & Blixt, T. (2008). The teaching of swimming based on a model derived from the causes of drowning. *International Journal of Aquatic Research and Education, 2*, 372–382. https://doi.org/10.25035/ijare.02.04.11

Tamplain, P., Brian, A., Webster, E. K., & Valentini, N.C. (2020). Assessment of motor development in childhood: Contemporary issues, considerations, and future directions. J*ournal of Motor Learning and Development, 8,* 391–409. https://Doi.org/10.1123/jmld.2018-0028

Vallerand, J. R. (1989). Ver une méthodologie de validation trans-culturelle de questionnaires psychologiques: Implications pour la recherche em langue française. *Canadian Journal of |Psychology, 30*(4), 23-38. https://Doi.org/10.1037/h0079856

Waltz, C.F., Strickland, O., & Lenz, E. R. (2010). *Measurement in nursing research*. Springer Publishing Company.

Wizer, R.T., Franken, M., & Castro, F. A. S. (2016). Within and between observer agreement in a protocol for aquatic skills assessment in children. *Revista Brasileira de Ciência e Movimento, 24*(1), 113-119. https://doi.org/10.18511/rbcm.v24i1.5835

World Health Organization (WHO, 2014). *Global Report on Drowning*. https://www.who.int/news-room/fact-sheets/detail/drowning