Motriz, Rio Claro, v.26, Issue02, 2020, e10200025 DOI: http://dx.doi.org/10.1590/s1980-6574202000010025

Original Article (short paper)

Inter- and intra-rater reliability of swimming teachers with different skill levels, in different conditions, evaluating front crawl arm movement in non-expert swimmers

Guilherme Tucher[1], Gustavo Ferreira Quintão[2], Nuno Domingos Garrido[3], Flávio Antônio de Souza Castro[4]

[1]*Universidade Federal do Rio de Janeiro, Escola de Educação Física e Desportos, Grupo de Pesquisa em Ciências dos Esportes Aquáticos, Rio de Janeiro, RJ, Brasil.*
[2]*Instituto Federal do Sudeste de Minas Gerais, Licenciatura em Educação Física, Rio Pomba, MG, Brasil.*
[3]*Universidade de Trás-os-Montes e Alto Douro, Departamento de Ciências do Desporto, Exercício & Saúde, Centro de Investigação em Ciências do Desporto, Saúde e Desenvolvimento Humano, Vila Real, Portugal.*
[4]*Universidade Federal do Rio Grande do Sul, Escola de Educação Física, Fisioterapia e Dança, Grupo de Pesquisa em Esportes Aquáticos, Porto Alegre, RS, Brasil.*

Abstract - Aim: This study aimed (i) to test the inter-rater reliability of swimming teachers, (ii) to test the swimming teacher's discussion effect on inter-rater reliability and (iii) to verify the intra-rater swimming teacher's reliability. **Method:** Twenty-one learning swimmers (14.1 ± 5.1 years old) performed two 25-m front crawl courses at a comfortable speed without breathing between the sixth and 20-m, and had their displacements captured on film. Three swimming teachers with different academic backgrounds and skills evaluated the swimmer right upper limb using a 20-items checklist. In the 1st-step, teachers assessed 20-items and in 2nd-step discussed their particular evaluating criteria - selecting five items considered as the most relevant. The inter- and intra-rater reliability were tested using the Fleiss Kappa Coefficient. **Results:** In the 1st-step substantial reliability was found for item three and in movement descriptor for items three and 20. Nearly perfect reliability was found in the movement descriptor for item 13. In 2nd-step, moderate reliability was found only in the movement descriptor for item 20. Only the most experienced evaluator showed substantial intra-rater reliability for items four and 10 and moderate for item 20. **Conclusion:** The proposed discussion method did not cause the expected effect on inter-rater reliability. The swimming teacher with a higher degree and swimming skills showed better intra-rater reliability. Some items and movement descriptors proposed at the 20-items checklist can be used in practical settings.

Keywords: swimming, sport, evaluation, pedagogy, teaching.

Introduction

At any swimming level learning is affected by the interaction of several components, including teacher/coach action[1-3]. Teachers, coaches, and scientists use different observational methods to evaluate and implement technique[1,4,5]. During swimming teaching, this evaluation usually takes a qualitative and subjective nature[1,6,7], where teachers/coaches with more knowledge and experience in the observational process tend to be more competent[7-9]. In an attempt to carry out a proper evaluation process, the literature suggests the analysis of the main components of movement[4] using checklists[6,7,10].

The inter and intra-rater reliability tests are common approaches that enable the satisfactory use of a checklist by different teachers and for the same teacher over time[6,10,11]. This reliability refers to the condition of reproducing similar measures on different occasions[11,12], being an important procedure in ensuring that the results do not have significant differences[6,8,10]. Furthermore, the inter-rater experiences should not influence the results, although the literature indicates that more experienced evaluators have better competences[8,10]. This is important because from this assessment that learning exercises will be proposed[5,10,11].

We recognize the importance of the results reported in the literature, but some checklists emphasize only movement errors, suggesting that there is only one rigid way to swim[7]. Indeed, the checklists do not (i) include the range of movements that can be performed by learners[7,10], and (ii) evaluators did not discuss their observation and evaluation criteria. Therefore, it is crucial to a movement checklist whose focus is based on comprehensive and accurate descriptors[4,13,14]. Little comprehensive criteria may not contemplate all the possibilities of movement, especially in the beginners. These two conditions make inter- and intra-rater reliability difficult and make the use of checklist unviable as a pedagogical tool[10]. Additionally, teachers use particular and subjective assessment criteria that differ for observation and evaluation moment[8,13]. Then, the discussion before evaluation can match observation points among evaluators and as far as we know, this has never been tested.

Accordingly, for the present study, three objectives were determined: (i) to test inter-rater reliability for a 20-item checklist and its movement descriptors; (ii) to test the effect of a discussion among teachers on selecting the most relevant items to assess and the own observation/evaluation criteria used, on the inter-rater reliability; and (iii) to check intra-rater reliability for the pre and post-intervention moments. We hypothesize that the inter-rater reliability would increase due to the discussion among teachers.

Materials and Methods

Three swimming teachers with different academic backgrounds and experience time (Table 1) were responsible for evaluating 21 swimming learners (average age 14.1 ± 5.1 years) from diverse swimming technical level. They did not have previous experiences with the checklist but were informed about how to use it. Non-expert swimmers were randomly selected from two swim clubs that have been training for at least 1 year. They could swim short distances of front crawl (up to 50-m) and some of them could swim backstroke. All participants received oral and written instructions about the study and participated voluntarily. Participants and their parents, when necessary, signed the Informed Consent Form. The Human Research Ethics Committee approved this study under number 2.486.694.

The right upper limb movement during the front crawl was assessed from a checklist and it was chosen because it is considered the most propulsive action[8,15].

Table 1 - Degree, years of swimming experience, and expertise area of the evaluators.

Evaluator	Degree	Experience (years)	Expertise
A	Specialization	12	Learning and fitness training
B	Specialization	4	Learning
C	Doctorate	18	Learning and competitive training

The complete checklist had 20-items (upper limb phases and sub-phases)[16,17] and different movement descriptors (MD), as shown in Chart 1. Due to the particularity of the movement, not all items had the same amount of MD. The MD is the possibility of the right upper limb movement during the front crawl.

Each swimmer performed two 25-m front crawl courses at a comfortable velocity without breathing between the sixth and 20-m to maintain standardization of the swimming technique and avoiding possible technical changes[18,19]. The swimming movements were captured on film in the sagittal (submerged and out of water) and frontal (submerged and out of water) planes in the central 10-m of the swimming pool, with a sports camera (Garmin, Virb Elite HD; frequency: 60Hz) and, when necessary, using a watertight box (Garmin). For recording out of the water the camera was stabilized by a tripod at 0.80 m high. To obtain underwater images the camera was kept at a depth of 0.45 m. The sagittal plane recording was performed about 5-m away from the swimmer, allowing the visualization of two-stroke cycles. The videos were later digitized using video editing software (Final Cut Pro).

Step 1 - Inter-rater evaluation

The evaluator received an e-mail link containing the swimmers' videos identified by letters and for three weeks evaluated them using the checklist presented in Chart 1. Each evaluator used his equipment, resources, criteria, and method for movement evaluation. The swimmers' videos could be viewed as many times as needed and without contact with the other evaluators.

Step 2 - Inter-rater evaluation after intervention meeting

Thirty days after the end of the first step, the evaluators and one of the researchers met for about 60 min to discuss the swimmers' evaluation process. Initially all items and MD were discussed.

After that, items 4, 10, 11, 15, and 20, were considered by evaluators as the most relevant for the assessment of upper limb movement in front crawl. Subsequently, these five items were again discussed among evaluators in order to align the observation and evaluation criteria. Movement interpretations from different observational planes and the possibility of slow-motion speed views were also discussed.

After the meeting, the evaluators received another e-mail link containing the swimmers' videos identified by letters in a different order from step 1. This time, they had two weeks to perform the new evaluation with five selected items. Again, the swimmers' videos could be viewed as many times as needed and without contact with the other evaluators.

Chart 1 - Items and movement descriptors (MD) of the right upper limb movement proposed for swimmer's assessment when performing the front crawl swim.

MD1	MD2	MD3	MD4
	Stroke Phase: Hand entry into the water		
	1. How does the hand enter the water?		
The hand enters in the water at the shoulder line, but very close to the head (short arm stroke).	The hand enters the water in or out of the shoulder line.	The hand enters the water at the shoulder line and in front of the head.	-
	2. How is hand positioned when entering the water?		
The hand enters the water with the palm facing inward.	The hand enters the water with the palm facing outwards.	The hand enters the water with the palm slightly facing down or out.	-
	3. Which part of the hand touches the water first?		
The fingertips are not the first part to touch the water.	The fingertips are the first part to touch the water.	-	-
	4. What is the direction of the hand when it enters the water?		
The hand is directed straight downward after touching the water.	The hand is directed straight inward after touching the water.	The hand is directed straight outward after touching the water.	The hand is directed straight ahead after touching the water.
	5. How is the arm positioned after the hand touches the water?		
The elbow is not fully extended after the hand touches the water.	The elbow is fully extended after the hand touches the water.	-	-
	6. How does the trunk moves after the hand touches the water?		
The trunk rotates excessively (> 45 °) after the hand touches the water.	The trunk does not rotate, or rotation is very discreet after the hand touches the water.	The trunk is continuously and slightly inclined (~ 45°) to the same side of the hand entering the water; however, there is no balanced rotation for both sides.	The trunk is continuously and slightly inclined (~ 45°) to the same side of the hand entering the water; there is a balanced rotation to both sides.
	7. How is head positioned throughout the entry and the glide of the hand in the water?		
The cervical spine is at hyperflexion (with the chin very close to the chest).	The cervical spine is hyperextended (with eyes facing forward).	The cervical spine is in a natural position and trunk extension (looking down or slightly forward).	-
	8. Which kind of movement the head performs throughout the entry and the glide of the hand in the water?		
It performs lateral movements.	It performs frontal movements.	It keeps unmoved.	-
	Stroke Phase: Catch		
	9. How is hand positioned throughout the catch phase?		
The palm is facing downwards (bottom of the pool).	The palm is facing inwards.	The palm is facing outwards.	The palm is facing backward (feet direction).
	10. How is elbow positioned at the catch phase?		
It is extended.	It is flexed and lies below the hand.	It is flexed and at the same level as the hand.	It is slightly flexed and higher than the hand.
	Stroke Phase: Pull and Push		
	11. How is hand positioning at the beginning of the pull until the end of the push phase?		
The palm is not facing backward throughout the phase.	Initially the palm is facing backward, but throughout the phase it turns inward.	Initially the palm is facing backward, but throughout the phase it turns outward.	The palm is facing backward initially and throughout the phase.
	12. How is the elbow angle at the end of the pull phase?		
It assumes an angle much lower than 90°.	It assumes an angle slightly greater than 90°.	It assumes an almost complete extension angle.	It assumes an angle of approximately 90°.
	13. How is the hand trajectory throughout the pull and push phase?		

(*continued*)

Chart 1 - continued

MD1	MD2	MD3	MD4
The hand crosses the midline of the trunk (to the opposite side).	The hand passes much laterally (out) from the midline of the trunk.	The hand passes a little laterally (out) from the midline of the trunk.	The hand passes close to the midline of the trunk but does not cross it.
14. How is hand positioned at the end of the push phase?			
The palm is facing inward.	The palm is facing outward.	The palm is facing backward.	-
15. How is the hand trajectory throughout the end of the push phase?			
The palm is directed downward.	The palm is directed outward.	The palm is directed inward.	The palm is directed backward and upward.
16. How is the elbow angle at the end of the push phase?			
It is not completely extended.	It is completely extended.	-	-
17. How is the hand speed throughout the subaquatic phase (pull and push)?			
It is performed at a constant speed.	It is performed with pauses.	It is performed at increasing speed.	-
18. How is the swimmer's displacement while performing front crawl stroke?			
Apparently, it is inefficient.	Apparently, it is slightly efficient.	Apparently, it is very efficient.	-
Stroke Phase: Recovery			
19. How is elbow movement throughout arm recovery?			
It remains more extended, being directed to the side and forward, by the side of the body.	It is flexed, directed straight forward, and slightly to the side of the body.	-	-
20. How is elbow angle throughout arm recovery?			
It is bent, low, close to water level or hand level.	It is at full or almost complete extension near the water level.	It is at full or almost complete extension, and far from the water level.	It is flexed, elevated, away from water, and above hand level.

Step 3 - Intra-rater evaluation

This step consisted of the evaluation of intra-rater reliability and was performed only for the five items selected in the second step of the study, considering the results of the first and second steps.

Statistical analysis

The MD used in the present study can be classified as nominal. The general level (item) of inter-rater (1st and 2nd step) and intra-rater (3rd step) reliability was tested using the Fleiss' kappa reliability coefficient (k). Using Fleiss' Kappa it was also possible to obtain the inter-rater reliability level for each of the MD. Fleiss' Kappa ranges from +1 (perfect reliability) to 0 (reliability equal to chance) and goes to -1 (complete unreliability)[20]. It is suggested that results between 0-0.2 show very little reliability, between 0.21-0.40 poor reliability, between 0.41-0.60 moderate reliability, between 0.61-0.80 substantial reliability and value above 0.80 as a nearly perfect reliability[21]. In all cases, a significance level of 95% was considered, and a reliability level above 0.41 (moderate) was considered relevant[20].

Results

Step 1 - Inter-rater evaluation

The inter-rater reliability for the 20-items indicates that there was moderate reliability for MD1 (The hand enters the water at the shoulder line, but very close to the head) and MD3 (The hand enters the water at the shoulder line and in front of the head.) of item 1 (how does hand enter in the water?); for MD1 (the hand is directed straight downward after touching the water) of item 4 (which does hand direction when it enters in the water?); on both MD and general item 5 (How is the arm positioned after the hand touches the water?); for MD1 (it performs lateral movements), MD3 (it keeps unmoved - when there is no breathing) and general item 8 (Which kind of movement the head performs throughout the entry and the glide of the hand in the water?); and for general item 13 (how is the hand trajectory throughout the pull and push phase?).

Substantial reliability has also been found for general item 3 (Which part of the hand touches the water first?) and their MD1 (the fingertips are not the first part to touch the water) and MD2 (the fingertips are the first part to touch the water); and for MD4 (it is flexed, elevated, away from water and above hand level) of item 20 (how is elbow angle throughout arm recovery?).

Nearly perfect reliability was found in MD1 (the hand crosses the midline of the trunk to the opposite side) of item 13 (how is the hand trajectory throughout the pull and push phase?); and MD3 (it is at full or almost complete extension, and far from the water level) of item 20 (how is elbow angle throughout arm recovery?). The complete result for the first step of the study is presented in Table 2.

Step 2 - Inter-rater evaluation after intervention meeting

The five items proposed in the second step of the study (after the discussion of observation and evaluation criteria) by evaluators showed moderate reliability only for MD2 (the elbow is at full or almost complete extension near the water level) and MD3 (the elbow is at full or almost complete extension, and far from the water level) of item 20 (how is elbow angle throughout arm recovery?). The complete result of the second step of the study is presented in Table 3.

Step 3 - Intra-rater evaluation

Evaluator A showed moderate internal reliability only in the evaluation of item 20 (how is elbow angle throughout arm recovery?). Evaluator B did not present at least one item with moderate reliability. On the other hand, evaluator C showed substantial reliability for items 4 (What is the direction of the hand when it enters the water?) and 10 (How is elbow positioned at the catch

Table 2 - Inter-rater reliability from the initial 20-items checklist and its movement descriptors (MD) of the high upper limb in front crawl proposed in the first step.

Item	General (sig)	MD1 (sig)	MD2 (sig)	MD3 (sig)	MD4 (sig)
1	0.38* (< 0.001)	0.55* (< 0.001)	0.21 (0.09)	0.41* (0.001)	-
2	0.26* (0.03)	-	0.26* (0.03)	0.26* (0.03)	-
3	0.67* (< 0.001)	0.67* (< 0.001)	0.67* (< 0.001)	-	-
4	0.34* (< 0.001)	0.48* (< 0.001)	- 0.05 (0.69)	- 0.06 (0.59)	0.37* (0.003)
5	0.48* (< 0.001)	0.48* (< 0.001)	0.48* (< 0.001)	-	-
6	0.16 (0.07)	-	0.22 (0.07)	0 (0.94)	0.19 (0.12)
7	- 0.14 (0.19)	- 0.16 (0.89)	- 0.14 (0.24)	- 0.16 (0.18)	-
8	0.48* (< 0.001)	0.57* (< 0.001)	- 0.03 (0.79)	0.52* (< 0.001)	-
9	0.02 (0.77)	0.07 (0.55)	- 0.01 (0.93)	- 0.68 (0.59)	0.04 (0.74)
10	0.13 (0.08)	0.23 (0.80)	0.16 (0.53)	0.33 (0.32)	0.58 (0.06)
11	0 (0.98)	0 (0.89)	0.31 (0.87)	0 (0.49)	0.60 (0.96)
12	0.07 (0.30)	0.11 (0.34)	0.16 (0.18)	0.04 (0.70)	0 (0.94)
13	0.42* (< 0.001)	1* (< 0.001)	0.27* (0.02)	0.04 (0.70)	0.32* (0.01)
14	0 (0.94)	0.02 (0.86)	- 0.05 (0.69)	0 (0.96)	-
15	0 (0.93)	-	- 0.86 (0.49)	- 0.01 (0.93)	0.02 (0.87)
16	0.02 (0.87)	0.02 (0.87)	0.02 (0.87)	-	-
17	0.26* (0.01)	0.38* (0.002)	- 0.03 (0.79)	0.21 (0.08)	- 0.16 (0.89)
18	0.06 (0.52)	- 0.08 (0.49)	- 0.01 (0.93)	0.22 (0.07)	-
19	- 0.02 (0.81)	- 0.02 (0.81)	- 0.02 (0.81)	-	-
20	0.29* (0.004)	0.53 (0.19)	0 (0.89)	0.80* (< 0.001)	0.62* (0.01)

*Significant at the 0.05
MD = movement descriptors.

Table 3 - Inter-rater reliability from the five items checklist and its movement descriptors (MD) of the high upper limb in front crawl proposed in the second step.

Item	General (sig)	MD1 (sig)	MD2 (sig)	MD3 (sig)	MD4 (sig)
4	0.31* (0.006)	0.28* (0.02)	-	- 0.33 (0.79)	0.39* (0.002)
10	0.19* (0.02)	0.17 (0.16)	0.30* (0.01)	0.35* (0.005)	0 (1.00)
11	0.16 (0.08)	- 0.06 (0.59)	0.17 (0.16)	0.30* (0.01)	0.19 (0.12)
15	- 0.12 (0.23)	-	- 0.06 (0.59)	- 0.14 (0.25)	- 0.12 (0.31)
20	0.31* (0.001)	0.17 (0.17)	0.48* (< 0.001)	0.44* (< 0.001)	0.40* (0.001)

*Significant at the 0.05;
MD = movement descriptors.

phase?), and moderate reliability for item 20 (how is elbow angle throughout arm recovery?). The result of intra-rater reliability is presented in Table 4.

Discussion

This study tested inter and intra-rater reliability of the upper limb technique in front crawl using a checklist. In this sense, the inter-rater reliability was initially evaluated for a 20-items checklist. After the intervention, to align the observation and evaluation criteria, the evaluators chose five items they considered most relevant and the inter-rater reliability was tested again. Finally, intra-rater reliability was tested for both evaluations. The study showed that only four items proposed on the initial checklist and 12 movement descriptors showed at least moderate reliability (Step 1). Moreover, none of the five items chosen, and only two of their movement descriptors had adequate inter-rater reliability (Step 2). Thus, the proposed intervention did not have the expected effect, indicating that teachers use different observation and evaluation criteria, certainly influenced by former experience. The most experienced evaluator was the one with the highest internal reliability. Even so, some items showed insufficient reliability.

The initial checklist was made up of 20-items. However, only three of them had moderate general inter-rater reliability and one substantial. This difficulty of reliability can happen because each evaluator has particular and subjective assessment criteria that differ for observation and evaluation moment[8,13]. Besides that, the way evaluators interpret the movement may influence their assessment - for example, does the evaluator refer to a more general (global body movement) or specific (concerned with a limb or segment) movement? Of these four items, three are easily observable outside the pool, where a swimming teacher usually makes his observations (items 3, 5, and 8) during class. Although item 13 refers to a movement made underwater, which could lead to difficulties in reliability, it is generally the focus of a teacher's observation because it refers to the underwater trajectory of the hand[8]. Therefore, these four items showed adequate inter-rater reliability and have practical implications. This finding may due to these items are commonly observed by teachers in swimming lessons.

Table 4 - Intra-rater reliability from the five items checklist and its movement descriptors (MD) of the high upper limb in front crawl proposed in the third step.

Evaluator	Item 4	Item 10	Item 11	Item 15	Item 20
A	0.32* (0.02)	0.03 (0.81)	0.11 (0.52)	0.31 (0.08)	0.57* (0.002)
B	0.39 (0.07)	- 0.01 (0.92)	- 0.20 (0.19)	- 0.02 (0.88)	0.14 (0.32)
C	0.62* (0.002)	0.72* (< 0.001)	0.28 (0.11)	- 0.08 (0.64)	0.48* (0.005)

*Significant at the 0.05.

Most items and MD that did not show at least moderate reliability are predominantly underwater movements and/or in combination with other more complex movements. Additionally, they can be less important during the learning process. Thus, the lack of experience in observing these movements daily may difficult their evaluation and, consequently, inter and intra-rater reliability levels[8,13]. Some movements are difficult to observe, either by the turbulence generated in the water or by the teacher's positioning (on the ledge or inside the water). For this reason, the teacher can evaluate these items by the consequence of movement and not by movement itself[1,22], with marked difference according to the evaluator's experience[8]. Nevertheless, despite the notion that a swimmer who presents a good quality of movement along with good performance must have a superior technique[1,8], sports performance is dependent on several variables[15,23]. Thus, a proper technique may lead to superior performance, but the opposite is not necessarily true[1]. Also, the focus of technical observation was on the right arm stroke and its associated movements. However, the overall movement of the swimming also depends on the action of the other upper and lower limbs[15,22].

Although some items did not show adequate inter-rater general reliability, some of their MD showed different behavior. While item 1 (How does the hand enter the water) showed poor general reliability (0.38), two of its MD showed moderate reliability (0.55 and 0.41 - respectively for MD1 and MD3). Therefore, in this case, the evaluators agreed to perceive the short arm stroke (MD1) and the arm that enters at the shoulder line and in the front of the head (MD3). Since observation and evaluation criteria differ between evaluators[8,13], the MD that showed the worst reliability was precisely the one in which comparison criteria became more subjective and particular. Consequently, the MD2 of item 1 (The hand enters the water in or out of the shoulder line.) showed poor reliability. This is because extreme cases of arm stroke crossing the midline of the body or very wide are easy to notice. Nevertheless, what is the limit between one arm stroke on the shoulder line and one stroke that slightly crosses the midline of the body? The teacher working at the learning level focuses his observation on gross errors and may not stick to the finer aspects of movement[10].

A similar understanding can be found for MD1 of item 4, MD1 of item 13 and MD3 and MD4 of item 20. It is more noticeable to watch the hand moves directly downward after entering the water (MD1 of item 4), as it is considered an extreme movement, than the subtle differences between directing the hand forward, a slightly inwards or outwards[10]. There are the same perception and reliability for the hand crossing the midline of the trunk (MD1 of item 13) than the other subtle variations of hand's path. In this case, for example, while the inter-rater reliability of the hand crossing the midline of the trunk in the underwater phase was perfect and significant (the best in the whole study), the inter-rater reliability of the hand passing away from the midline of the body (MD2) and the hand passing near the midline of the trunk (MD4) was only poor. This may also indicate that inter-rater reliability for movements with less execution variance is more difficult to perceive in the same way, especially for the less experienced[8,10]. In some cases this can cause difficulties in perceiving movement error bringing consequences for learning[10,13].

It can be thought that it is easier to find appropriate inter-rater reliability when MDs are mutually exclusive. In fact, this happened at items 3 (Which part of the hand touches the water first?) and at item 5 (How is the arm positioned after the hand touches the water?). However, not at item 16 (how is elbow angle at the end of the push phase?) and item 19 (how is elbow movement throughout arm recovery?). It may seem easy, at item 16, to indicate only whether there is an elbow extension at the end of the pushing phase. However, the analysis involves the correct understanding of (i) the end of this phase, (ii) the hand speed, and (iii) whether the elbow extension is appropriate. Thus, great movements can be confused with their performance consequence (displacement efficiency)[1,10,22]. Item 19 becomes complex because it subjectively involves a kinematic and temporal analysis. Between a flexed and extended elbow may be an angle that raises doubts. In addition, the movement may start in one way, and along its path assume another behavior.

Importantly, while technical analysis is imperative for teachers, coaches, and scientists, their interests and tools are different[1]. This study was concerned with the swimming teachers and their particularities in the teaching process. Of the five items selected by the evaluators as most relevant, four comprised underwater movements and one out of water. The evaluators' focus seems to be on hand entry into the water, arm positioning at the beginning of the propulsive phase, and the direction and trajectory of the hand in this propulsive action. The relevance of these items was also highlighted in another study[8].

Despite the recognition of the importance of these five items for an adequate technique, the observation and evaluation criteria were very different between evaluators. In fact, the inter-rater reliability for item 10 (How is elbow positioned at the catch phase), item 11 (how is hand positioned at the beginning of the pull until the end of the push phase?) and item 15 (how is the hand trajectory throughout the end of the push phase?) were small and not significant[8,10]. This seems to be a problem for the learning process because with different assessment swimmers would take different feedbacks[24,25].

It was hypothesized that the inter-rater reliability for the five selected items would increase due to the implemented intervention. However, the overall reliability rating for the five items did not change. Nevertheless, MD2 (it is flexed and lies below the hand) and MD3 (it is flexed and at the same level as the hand) of item 10 (How is elbow positioned at the catch phase), MD3 (initially the palm is facing backward, but throughout the phase, it turns outward) of item 11 (how is hand positioned at the beginning of the pull until the end of the push phase?) and MD2 (it is at full or almost complete extension near the water level) of item 20 (how is elbow angle throughout arm recovery?) showed an improvement in their reliability rating, although they are still low.

The MD3 (it is at full or almost complete extension, and far from the water level) and MD4 (it is flexed, elevated, away from water and above hand level) of item 20 (how is elbow angle throughout arm recovery?) had their reliability ratings reduced, respectively, from practically perfect and moderate, to moderate and poor. The same happened with MD1 (the hand is directed straight downward after touching the water) of item 4 (What is the direction of the hand when it enters the water?), which had its reliability rating changed from moderate to poor. Item 15 (how is the hand trajectory throughout the end of the push phase?) and its MD continued to show random reliability or even marked unreliability. These results indicate that some observation and evaluation points could be slightly adjusted among the evaluators due to the intervention implemented, but not enough for the item to present substantial or practically perfect reliability. Since it is an important item, an effort should be made to approximate the observation and evaluation strategy used[8]. Maybe it can occur with a greater period of discussion between teachers.

Intra-rater reliability was lower for the evaluator with less experience in swimming teaching. This may indicate that the evaluator is still forming his observation, comparison, and analysis references[8,13]. However, even evaluator B, with more than 10 years of teaching experience, presented only one item with moderate reliability. Anyway, in these two cases, it is worth noting that there was an intervention between the two moments of evaluation, which may have been a factor that led to some behavioral change and the absence of more expressive reliability. Explicitly, evaluators may have used new observation and analysis criteria that contributed to these results[8]. Usually the teacher compares the swimmer's movement with a standard model that he has as a reference and then makes his decisions[1,10]. Thus, it seems that the standard models created by evaluators are different, and, for some movements, their internal standard models of comparison are not well defined, making even their internal reliability difficult[8].

The evaluator with the longest experience and academic background was the one who presented substantial intra-rater reliability for two items (4 and 10) and moderate intra-rater reliability for one (20). More experienced evaluators were able to differentiate swimmers' performance levels more accurately than those less experienced[8]. On the other hand, it is not known whether the intervention strengthened or was unable to modify some of its observation and evaluation criteria in light of its pre-established concepts. Regardless, it should be noted that even in the most experienced evaluator, items 11 and 15 presented an unsatisfactory classification of reliability, which had already happened in the inter-rater reliability. Therefore, one must reinforce that the perception and evaluation of the movement and direction of the hand during the underwater actions of the stroke are difficult even with video. These points should be addressed, as they are extremely important in swimming technique and may lead to inadequate feedback by swimming teachers[8,24,25].

The findings of this study have to be seen in the light of some limitations. The first is how teachers use the computer and its resources to better observe and understand the movement. At step 1 they assessed the swimmers from personal criteria and at step 2 some observational tips were shown. Nevertheless, they used the computer as their personal preferences. The second limitation concerns the observational focus[14]. We did not evaluate the front crawl, but just the right upper limb movement. To enumerate, breath and arm coordination were not evaluated, and the outcome could be different. Breathing, for instance, may cause technical changes[18,19] and it should be one specific observational focus, as proposed in literature[4,14].

Conclusion

Four of the 20-item checklist had moderate and substantial inter-rater reliability. In addition, 12 movement descriptors had at least moderate inter-rater reliability. These items and movement descriptors proposed at the 20-items checklist can be used in practical settings. Although five items were considered as the most relevant for the swimmer's technical evaluation, only two movement descriptors showed adequate inter-rater reliability. Therefore, the novelty proposed discussion method for improving inter-rater reliability did not cause the expected effect. Hence, it was unable to modify some of the teacher observation and evaluation criteria in light of its pre-established concepts. Moreover, the intra-rater reliability can be influenced by the evaluator degree and experience. The better intra-rater reliability was found with a higher degree and experience swimming teaching.

References

1. Lees A. Technique analysis in sports: a critical review. J Sports Sci. 2002; 20(10): 813-28.
2. Gonzalez DH, Chiviacowsky S. Relatedness support enhances motor learning. Psychol Res. 2018; 82(3): 439-47.
3. Wulf G, Shea C, Lewthwaite R. Motor skill learning and performance: a review of influential factors. Medical Education. 2010; 44(1): 75-84.
4. Federolf P, Reid R, Gilgien M, Haugen P, Smith G. The application of principal component analysis to quantify technique in sports. Scand J Med Sci Sports. 2014; 24(3): 491-9.
5. Potdevin F, Vors O, Huchez A, Lamour M, Davids K, Schnitzler C. How can video feedback be used in physical education to support novice learning in gymnastics? Effects on motor learning, self-assessment, and motivation. Phys Educ Sport Pedagogy. 2018; 23(6): 559-74.
6. Wizer R, Franken M, Castro FAdS. Concordância intra e inter-observador de protocolo de avaliação de habilidades aquáticas de crianças. R. Bras. Ci. e Mov. 2016; 24(1): 101-7.
7. Madureira F, Gollegã DG, Rodrigues HF, de Oliveira TAC, Dubas JP, Freudenheim AM. Validação de um instrumento para avaliação qualitativa do nado “Crawl”. Rev Bras Educ Fís Esporte. 2008; 22(4): 273-84.
8. Zamparo P, Carrara S, Cesari P. Movement evaluation of front crawl swimming: technical skill versus aesthetic quality. PloS One. 2017; 12(9): 1-12.
9. Rosado A, Virtuoso L, Mesquita I. Relação entre as competências de diagnóstico de erros das habilidades técnicas e a prescrição pedagógica no voleibol. Rev Bras Educ Fís Esporte. 2004; 18(2): 151-7.
10. Soares S, Fernandes R, Carmo C, Santos Silva J, Vilas-Boas J. Avaliação qualitativa da técnica em natação: apreciação da consistência de resultados produzidos por avaliadores com experiência e formação similares. Rev. Port. Ciênc. Desporto. 2001; 1(3): 22-32.
11. Nadeau L, Richard J-F, Godbout P. The validity and reliability of a performance assessment procedure in ice hockey. Phys Educ Sport Pedagogy. 2008; 13(1): 65-83.
12. Currell K, Jeukendrup AE. Validity, reliability, and sensitivity of measures of sporting performance. Sports Med. 2008; 38(4): 297-316.
13. Mendes R, Clemente F, Rocha R, Damásio A. Observação como instrumento no processo de avaliação em educação fisica. Exedra Journal. 2012; 6(57-69).
14. Knudson D, Morrison CS. Análise qualitativa do movimento humano. 1 ed. São Paulo, SP: Manole, 2001.
15. Silveira RP, de Souza Castro FA, Figueiredo P, Vilas-Boas JP, Zamparo P. The effects of leg kick on swimming speed and arm-stroke efficiency in the front crawl. Int J Sports Physiol Perform. 2017; 12(6): 728-35.
16. Chollet D, Chalies S, Chatard J. A new index of coordination for the crawl: description and usefulness. Int J Sports Med. 2000; 21(1): 54-9.
17. Maglischo EW. Nadando o mais rápido possível. 3 ed. Barueri, SP: Manole, 2010.
18. Lerda R, Cardelli C. Breathing, and propelling in crawl as a function of skill and swim velocity. Int J Sports Med. 2003; 24(01): 75-80.
19. Psycharakis SG, McCabe C. Shoulder and hip roll differences between breathing and non-breathing conditions in front crawl swimming. J Biomech. 2011; 44(9): 1752-6.
20. Field A. Descobrindo a estatística usando o SPSS. 1 ed. Porto Alegre: Artmed, 2009.

21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977: 159-74.
22. Madureira F, Bastos FH, Corrêa UC, Rogel T, Freudenheim AM. Assessment of beginners' front-crawl stroke efficiency. Percept Mot Skills. 2012; 115(1): 300-8.
23. Costa MJ, Bragada JA, Mejias JE, Louro H, Marinho DA, Silva AJ, et al. Tracking the performance, energetics, and biomechanics of international versus national level swimmers during a competitive season. Eur J Appl Physiol. 2012; 112(3): 811-20.
24. Wulf G. Attentional focus and motor learning: a review of 15 years. International Review of Sport and Exercise Psychology. 2013; 6(1): 77-104.doi: 10.1080/1750984X.2012.723728
25. Marques-Dahi MTdSP, Bastos FH, de Araujo UO, Walter C, Freudenheim AM. Verbal instructions on learning the front-crawl: emphasizing a single component or the interaction between components? Human Movement. 2016; 17(2): 80-6.

Corresponding author
Guilherme Tucher Universidade Federal do Rio de Janeiro, Escola de Educação Física e Desportos, Departamento de Corridas.Postal address: Avenida Carlos Chagas Filho, 540, Sala A-225, 21941-599, Cidade Universitária, Rio de Janeiro, RJ. Brazil.
E-mail: g.tucher@eefd.ufrj.br.

Manuscript received on February 3, 2020
Manuscript accepted on April 20, 2020