Manchester Metropolitan University

Nicholson, VP and Spathis, JG and Hogarth, LW and Connick, MJ and Beckman, EM and Tweedy, SM and Payton, CJ and Burkett, BJ (2018) *Establishing the reliability of a novel battery of range of motion tests to enable evidence-based classification in Para Swimming.* Physical Therapy in Sport, 32. pp. 34-41. ISSN 1466-853X

Publisher: Elsevier

DOI: https://doi.org/10.1016/j.ptsp.2018.04.021

ESTABLISHING THE RELIABILITY OF A NOVEL BATTERY OF RANGE OF MOTION TESTS TO ENABLE EVIDENCE-BASED CLASSIFICATION IN PARA SWIMMING

Vaughan Nicholson[1], Jemima Spathis[2], Luke Hogarth[3], Mark Connick[4], Emma Beckman[4], Sean Tweedy[4], Carl Payton[5], Brendan Burkett[3]

Affiliations:

[1] School of Physiotherapy, Australian Catholic University, Brisbane, Australia

[2] School of Exercise Science, Australian Catholic University, Brisbane, Australia

[3] School of Health and Sport Sciences, University of the Sunshine Coast, Sippy Downs, Australia.

[4] School of Human Movement Studies, University of Queensland, St Lucia, Queensland, Australia.

[5] HEAL Research Centre, Manchester Metropolitan University, Crewe, United Kingdom.

Corresponding author: Vaughan Nicholson, School of Physiotherapy, Australian Catholic University, PO Box 456 Virginia Queensland 4014 (Australia). Phone: +61 7 36237687 Email: vaughan.nicholson@acu.edu.au

This research was supported by the International Paralympic Committee (IPC), UK Sport, and Exercise & Sports Science Australia (Applied Sports Science Research Grant). Mark Connick, Emma Beckman and Sean Tweedy are members of the IPC Classification Research and Development Centre (Physical Impairments), which is supported by the International Paralympic Committee. The authors would like to acknowledge Jayden Lowrie, Samantha Yardy-Phelan and Ana Maia for their assistance in data collection.

Abstract

Objectives: To evaluate the reliability of swimming-specific range of movement tests developed in order to permit evidenced-based classification in the sport of para swimming.

Design: Test-retest intra- and inter-examiner reliability.

Setting: International Swimming training camps and university exercise science departments.

Participants: 42 non-disabled participants (mean age 23.2 years) and 24 Para swimmers (mean age 28.5 years).

Main Outcome Measures: Intra- and inter-examiner reliability of a battery of novel active range of motion tests.

Results: Good to excellent intra-examiner reliability was found for the majority (32/34) of tests in non-disabled participants (ICC = 0.85-0.98). SEM values ranged from 1.18° to 6.11°. Similarly, good to excellent inter-examiner reliability was found for the majority (35/42) of tests in non-disabled participants (ICC = 0.85-0.98). SEM values range from 0.73° to 6.52°. Para swimmers exhibited significantly reduced range of motion compared to non-disabled participants.

Conclusions: The large majority of ROM tests included in this novel battery were reliable both within and between examiners in non-disabled participants. The tests were found to differentiate between non-disabled participants and Para swimmers with hypertonia or impaired muscle power.

Key words:

Swimming; Shoulder; Hip; Inclinometer

INTRODUCTION

Paralympic classification systems aim to promote participation in sport by people with disabilities by minimizing the impact impairment has on the outcome of competition (Tweedy & Vanlandewijck, 2011). Classification systems which achieve this aim will ensure that successful athletes are not simply those with the least impairment, but will be those that have the most advantageous combination of physiological and/or psychological attributes and who have trained them to best effect (Tweedy & Vanlandewijck, 2011). In 2007, the IPC Classification Code mandated the development of evidence-based systems of classification in Paralympic sport which are informed by scientific research (IPC, 2007).

The Paralympic games are the largest organized sporting event for athletes with disabilities with 164 participating countries and more than 1.8 million tickets sold at the 2016 Rio Paralympic Games (IPC, 2016). In Para Swimming, there are eight physical impairments comprising; impaired muscle strength, impaired passive range of movement (ROM), limb deficiency, leg length difference, short stature, hypertonia, ataxia, and athetosis (IPC, 2007). Eligible swimmers compete across ten classes in freestyle, backstroke and butterfly events (S1-S10) and nine classes in breaststroke events (SB1-SB9). The current classification process involves two-steps: 1) evaluation of impairment via a bench test (Dummer, 1999) and 2) a technical assessment known as a water test. The bench test involves evaluation of joint range of motion, which may be critical to determining an athlete's class in Para swimming. Impaired ROM is defined as a reduction in one or more joints which is reduced permanently (IPC, 2007). Impaired ROM from health conditions such as cerebral palsy, resulting in spasticity or contracture, for example are currently evaluated using a passive joint-by-joint assessment conducted by a trained classifier using a goniometer which is then scored against a passive functional ROM on a zero to five scale (IPC, 2017). The water test involves allocation of points based on an athlete's ability to perform key water skills specific to swimming such as a dive start and push-off when turning. The classifier then totals the points obtained from the bench test and the water test using established criteria and expert opinion to determine a class (IPC, 2017).

For classification, these measures of impaired ROM in the bench test have advantages in that they use widely utilized methods known to clinicians and the equipment is inexpensive and available world-wide. However, there are several limitations of these measures, which make them unsuitable methods of ROM assessment for classification. Firstly, the current system utilizes a ratio scaled measure with a goniometer (Tweedy & Vanlandewijck, 2011) but then converts the outcome to a point scale (zero to five) based on relative range of motion rather than allocating absolute values. As such, an athlete that scores four points on a ROM measure does not necessarily have twice the ROM as someone that scores two points. Weak relationships have been found between non-ratio scale measures currently used in classification and sports performance (Beckman, Connick, & Tweedy, 2016). Secondly, ROM is currently assessed using passive ROM techniques, where the classifier moves the athlete's joints through a range applying external forces to the body, which has questionable repeatability compared to active ROM techniques (Boon & Smith, 2000; Cools, et al., 2014; De Winter, et al., 2004; Muir, Corea, & Beaupre, 2010). Thirdly, the reliability of the majority of ROM tests currently included in the classification system are not known or their reliability is difficult to determine due to significant variability in the positioning of the participant and the equipment selected (i.e. inclinometer, goniometer, visual assessment or a combination) (van de Pol, van Trijffel, & Lucas, 2010). Fourthly, ROM is assessed using a joint-by-joint method which is not parsimonious, is time intensive and little evidence exists to show the impact of individual joint ROM on swimming performance. For example, only weak correlations exist between current ROM classification measures and swimming propulsion and joint kinematics (Evershed, Frazer, Mellifont, & Burkett, 2012). Additionally, measures of impairment should only assess body structures that will impact on performance in body positions relevant to sports performance (Tweedy & Vanlandewijck, 2011) and that can be achieved by all eligible athletes regardless of impairment type or severity. The current ROM tests do not necessarily assess ROM in positions relevant to swimming, and neglect certain swimming specific positions such as streamline and prone shoulder flexion. These limitations threaten the validity of the classification system and can result in failure to delineate performance between some classes and disadvantaging athletes with certain types of physical impairment within classes (Evershed, et al., 2012; Howe & Jones, 2006; Oh, Burkett, Osborough, Formosa, & Payton, 2013).

Given the limitations of current passive ROM tests and in order to permit evidence-based methods of classification, the development of a battery of novel ROM measures for swimming is required (Tweedy, Beckman, & Connick, 2014). The characteristics of the battery are required to comply with the IPC Position Stand on Classification which states that tests of impairment be objective, reliable, precise, ratio-scaled and valid (Tweedy, et al., 2014; Tweedy & Vanlandewijck, 2011). Additionally, the tests of impairment should be impairment specific, as resistant to training as possible, comprehensive by addressing movement relevant to the sport, and parsimonious.

The aim of the study was to enable evidence-based classification in Para swimming by establishing the reliability and preliminary normative values for a battery of swimming-specific range of motion tests by: 1) evaluating the intra-examiner and inter-examiner reliability of each novel ROM measure in non-disabled participants; 2) providing preliminary normative values for a novel ROM assessment battery in non-disabled participants and a group of trained Para swimmers; and 3) determining if differences exist for the novel ROM battery outcomes between non-disabled participants and Para swimmers.

METHODS

Participants

Two groups of participants were tested: Group 1 comprised 42 (20 males, 22 females) non-disabled physically active, injury free individuals engaging in at least 90 minutes of moderate physical activity per week. These participants were recruited from two Australian Universities. Group 2 comprised 24 (17 males, 7 females) elite Para swimmers who were nationally or internationally classified (classes S1-S8) and were undertaking planned training regimes and competing at national or international level. Group 2 included athletes with spinal cord injury (SCI), polio, cerebral palsy (CP) or acquired brain injury (ABI) - conditions that can result in one of, or a combination of impaired ROM, impaired strength or impaired coordination. These Para swimmers were categorized into two subgroups: those with hypertonia associated with CP and ABI (n=11, 9 males, 2 females); and those with impaired muscle power resulting from SCI or polio (n=13, 8 males, 5 females). While both subgroups may demonstrate some impairment in ROM, each subgroup has different ROM patterns due to the nature of their health condition. Athletes were from England, Spain, Italy and Czech Republic.

Study Design and Procedures

Data collection was conducted by three PhD trained staff with professional qualifications in the movement sciences with experience working with elite athletes with disabilities. All reliability testing was undertaken within University Exercise Science facilities. Assessment of Para swimmers was undertaken at various training facilities within Europe. All participants provided written informed consent and the study was approved by the Human Research Ethics Committees of the lead author's institution.

Non-disabled participants and Para swimmers completed a baseline questionnaire regarding demographics, training (number and types of sessions per week) and injury history. Anthropometric measures (standing height (m), body mass (kg), body mass index (BMI)) were also assessed prior to ROM assessment. Range of motion (ROM) testing was conducted with an Acumar Digital Inclinometer (Lafayette Instrument Co. Lafayette, IN), which was zeroed and aligned with the appropriate reference point (horizontal or vertical) before each test. A universal goniometer (Baseline Evaluation Instruments, White Plains, NY) was also used to compare reliability and ease of administration for elbow flexion and extension (inter-examiner study). Trunk functional reach measures were obtained with a supported fixed tape measure.

Measures from the 42 non-disabled participants were used to establish preliminary normative values. Fifteen of these non-disabled participants also took part in intra-examiner reliability testing while a further 16 took part in inter-examiner testing. Para swimmers completed the ROM test battery on one occasion to provide preliminary normative data for the ROM tests and to compare whether differences existed between subgroups of Para swimmers and between Para swimmers and non-disabled participants. Intra-examiner reliability: fifteen non-disabled participants completed two, one-hour testing sessions. Participants were tested by a single examiner on two occasions with at least one day between testing sessions. Two trials of each test were conducted on each participant following a practice trial, and the average of the results used for analysis.

Inter-examiner reliability testing: sixteen non-disabled participants completed a single two-hour testing session and were tested independently by two examiners for all ROM tests within a single session. Two trials of each test were conducted on each participant following a practice trial, and the average of the results used for analysis. Testing followed the same standardized order for each participant for both the intra- and inter-examiner reliability protocols.

Tests of range of motion (ROM)

Active range of motion was assessed via a battery of 10 tests designed to measure shoulder, elbow, hip, knee, ankle and trunk motion relevant to S-class swimming events (freestyle, backstroke and butterfly strokes). These 10 tests produced over 30 measures as some tests assessed multiple joints. All measures were demonstrated by the examiner before being performed by the participant. The following landmarks were located and marked with a semi-permanent pen prior to ROM assessment to provide consistency of inclinometer placement: lateral aspect of acromion process; lateral epicondyle; ulna and radial styloid processes; greater trochanter; lateral knee joint line; tibial tuberosity; lateral malleolus; anterior aspect of talus; dorsal aspect of 2^{nd} metatarsophalangeal (MTP) joint. Mid-points between adjacent landmarks were measured and marked, then used to place the digital inclinometer. For example, the mid-point between the lateral aspect of the acromion process and lateral epicondyle was used as the upper arm landmark to measure shoulder abduction. Detailed descriptions for all tests can be found in supplementary material (supplementary Table 1).

i) Bilateral Shoulder flexion (streamline)

ii) Bilateral Shoulder abduction (streamline)

iii) Elbow flexion and extension

iv) Lower-limb streamline (hip, knee and ankle extension)

v) Hip and knee flexion

vi) Shoulder internal and external rotation

vii) Prone shoulder extension (in elbow flexion and extension)

viii) Prone shoulder horizontal abduction

ix) Prone shoulder flexion

x) Trunk functional reach (forward, backward, sideways)

Data analysis

Data were analyzed using IBM SPSS 22 for Windows. All variables were examined for normality using the Shapiro-Wilk test. A two-way mixed model intra-class coefficient ($ICC_{3,2}$) was used to determine reliability between session one and two for intra-examiner reliability, and between examiners for inter-examiner reliability (Shrout & Fleiss, 1979) for non-disabled participants. Absolute agreement between sessions for intra-examiner and between examiners for inter-examiner reliability was based on the mean of two values from each session and examiner, respectively. Good to excellent reliability was defined a priori as an ICC > 0.75, moderate reliability defined as an ICC 0.5-0.75 and poor reliability defined as an ICC < 0.5 (Portney & Watkins, 2009). ICC values may be high despite poor trial to trial consistency if the inter-subject variability is too high (Weir, 2005), to negate this issue, the standard error of measurement (SEM) [$SEM = \text{Average Standard deviation} \times \sqrt{(1 - ICC)}$] was also calculated as this is not affected by inter-subject variability.

Paired-samples T-tests were conducted to identify any differences between testing sessions or between examiners. One-way ANOVA was used to identify differences in ROM between non-disabled participants, Para swimmers with hypertonia and Para swimmers with impaired muscle power. Post-hoc comparisons using Tukey HSD were applied when a difference between means was identified within the ANOVA. Significance was set at alpha < 0.05.

RESULTS

Intra-examiner reliability

Fifteen non-disabled participants aged 21.9 (± 3.4) years took part in the intra-examiner reliability study. Each participant was tested by the same examiner with each test session separated by 4.1 (± 2.7) days. The majority of tests produced good to excellent (>0.75) ICC absolute agreement values (Table 2). There were three exceptions where moderate reliability values were obtained: bilateral shoulder abduction ($ICC_{3,2} = 0.73$), functional reach forwards ($ICC_{3,2} = 0.66$) and functional reach right ($ICC_{3,2} = 0.68$). SEM values ranged from 1.18° to 6.11°.

There were no significant differences between sessions for any measures except elbow extension left ($t = 2.32$, $p = 0.04$) with goniometer, lower-limb streamline left ankle angle ($t = 3.93$, $p < 0.01$), and knee flexion left with goniometer ($t = -2.22$, $p = 0.04$) (Table 2).

Inter-examiner reliability

For the inter-examiner reliability study, 16 non-disabled participants aged 25.1 (±5.1) years took part in the study. All tests produced good to excellent (>0.75) ICC absolute agreement values (Table 3) except for six measures that produced moderate reliability: elbow extension right with inclinometer ($ICC_{3,2} = 0.75$); right elbow flexion with goniometer ($ICC_{3,2} = 0.67$) and right elbow extension ($ICC_{3,2} = 0.73$) with goniometer; knee flexion right with goniometer ($ICC_{3,2} = 0.72$), functional reach forwards ($ICC_{3,2} = 0.73$) and right ($ICC_{3,2} = 0.62$) together with one measure that produced poor reliability: knee flexion left with goniometer ($ICC_{3,2} = 0.21$). SEM values ranged from 0.73° to 6.52°.

There were significant differences between examiners for right elbow flexion ($t = 2.41$, $p = 0.03$) and extension ($t = -2.88$, $p = 0.01$) with goniometer, left elbow flexion with inclinometer ($t = -3.96$, $p < 0.01$) and goniometer ($t = 5.42$, $p < 0.01$), hip flexion left ($t = 2.69$, $p = 0.02$), knee flexion left with goniometer ($t = -2.25$, $p = 0.04$), shoulder internal rotation left ($t = -3.10$, $p = <0.01$) and right ($t = -2.71$, $p = 0.02$), (Table 3).

Preliminary normative values

There were significant differences between non-disabled participants and Para swimmers for the majority of ROM tests (Table 4). Significant differences were also found between Para swimmers with hypertonia and Para swimmers with impaired muscle power for certain measures at the trunk, hip and knee (Table 4). Participant characteristics are presented in Table 1 and preliminary normative values for each group are presented in Table 4.

DISCUSSION

The aim of this research was to enable evidence-based classification in Para swimming by establishing the reliability and normative values for a battery of swimming-specific range of motion (ROM) tests for swimmers. The results presented herein demonstrate that the majority of active ROM assessments were reliable in non-disabled participants, and Para swimmers had significantly less ROM than non-disabled participants. This finding addresses the key guidelines for the international classification system, that is the measures of impairment for the purposes of classification should be ratio scaled, reliable, precise and comprehensive by addressing movement relevant to the sport (Tweedy & Vanlandewijck, 2011).

The key finding of this study is that the majority of ROM assessments used in this novel test battery showed good to excellent levels of reliability in non-disabled participants. Establishing that the measures are reliable within non-disabled participants is an essential step towards developing evidence-based classification systems (Beckman & Tweedy, 2009). One of the important differences between this novel test battery and the ROM assessments currently employed for Para swimming classification is that this test battery assessed active ROM while the Para swimming classification tests measure passive ROM. Active range of motion measurements are deemed to be more reliable than passive measures, particularly in the shoulder (Boon & Smith, 2000; Cools, et al., 2014; De Winter, et al., 2004; Muir, et al., 2010) which is of particular interest in Para swimmers. When comparing our results to previous ROM reliability studies only a few measures can be accurately scrutinized, as although some of the tests within this battery are used routinely within swimming screenings and clinical assessments (Blanch, 2004), there is a paucity of literature examining their reliability. Shoulder internal rotation and external rotation are two exceptions as they are measures that have received repeated attention within the literature (Cools, et al., 2014; Furness, Johnstone, Hing, Abbott, & Climstein, 2015; Riemann, Witt, & Davies, 2011; Walker, et al., 2016). The intra-examiner values in this study (>0.80) for both internal and external rotation were superior to some (Awan, Smith, & Boon, 2002; Walker, et al., 2016) but slightly lower than others that have reported ICC values of greater than 0.9 (Cools, et al., 2014; Furness, et al., 2015) using an inclinometer. The good to excellent reliability values found for inter-examiner assessment are similar to previous studies that have also found good to excellent reliability values both in swimmers (Riemann, et al., 2011; Walker, et al., 2016) and non-swimmers (Cools, et al., 2014; Furness, et al., 2015). The generally low levels of SEM in this study indicate consistency across testing sessions and between examiners for this test battery and compare favorably to previous reliability studies that have reported SEM values of 2-5° for active shoulder movements (Kolber, Vega Jr, Widmayer, & Cheng, 2011; Walker, et al., 2016).

The ability to maintain optimal lower limb and trunk positions is important for swim performance (Chatard, Lavoie, Bourgoin, & Lacour, 1990; Daly, Malone, Smith, Vanlandewijck, & Steadward, 2001; Zamparo, Gatta, Pendergast, & Capelli, 2009) but there is limited literature assessing the reliability of lower limb measures in swimmers or young athletes as the majority of lower limb reliability studies have been conducted in patient populations (van Trijffel, van de Pol, Oostendorp, & Lucas, 2010). All lower limb tests in this study that were measured with an inclinometer produced good to excellent levels of reliability. The lower limb reliability values and SEMs established in our study for active ROM are superior to the majority of lower limb passive ROM measures previously published (Currier, et al., 2007; van Trijffel, et al., 2010). For example, our results for lower limb streamline ankle plantar flexion achieved excellent ICC values which is in contrast to previous studies assessing plantar flexion that have reported poor to moderate levels of inter-examiner reliability when assessed either actively (Youdas, Bogard, & Suman, 1993) or passively (Elveru, Rothstein, & Lamb, 1988) with a goniometer. The superior results in our study are likely due to the use of an inclinometer rather than a goniometer.

The choice of measuring instrument is an important consideration in ROM assessment and consequently classification. The majority of tests that did not achieve good to excellent reliability in this test battery were measured using a goniometer. Some of these measures such as knee flexion and elbow flexion, were more reliable when obtained with an inclinometer, suggesting that the inclinometer be the preferred method of measuring these movements, which is in contrast to current Para swimming protocols that still use goniometers for ROM measures (IPC, 2017). Only one measure across the test battery, left knee flexion with goniometer, achieved a poor level of reliability (Portney & Watkins, 2009), suggesting that it may not be suitable for classification purposes. The other tests that did not consistently achieve good to excellent reliability were functional reach forwards and sideways. The reduced reliability in these trunk related measures was most likely due to small variations in participants' sitting posture and difficulty maintaining the end position long enough to obtain an accurate measurement. Despite these limited reliability concerns, the tests were easily administered in non-disabled participants without the need for expensive equipment – a requirement for any potential classification battery as the tests should facilitate international dissemination and implementation (Tweedy & Vanlandewijck, 2011).

While the results indicate that the majority of tests used in this novel ROM battery are reliable, classification methods need to be comprehensive by addressing movement relevant to the sport (Tweedy & Vanlandewijck, 2011). The current classification system assesses passive shoulder extension (in elbow extension) in prone but it does not assess active shoulder flexion or horizontal abduction range of motion in prone. The battery of tests used in this study incorporated two measures (tests vii) of shoulder extension (in both elbow flexion and extension) together with horizontal abduction (test viii) as these tests replicate the arm positions during the recovery phase of freestyle (Pink, Perry, Browne, Scovazzo, & Kerrigan, 1991). Prone shoulder flexion was also incorporated (test ix) to replicate the entry and early pull phases of freestyle (Payton, Bartlett, Baltzopoulos, & Coombs, 1999; Pink, et al., 1991). This unilateral shoulder flexion assessment is distinct from the bilateral streamline position assessed in sitting (test i), as a swimmer's upper limb streamline position is limited by the least mobile shoulder, but a swimmer's hand entry position or early pull position can potentially vary substantially between sides (Evershed, et al., 2012). It is therefore important that the ROM of both sides are captured independently otherwise an athlete with a relatively poor streamline position due to limitations on one side will have an erroneous representation of their actual ability to propel themselves with their less impaired side. Another important addition to this test battery was lower limb streamline (test iv) which assesses ankle plantar flexion with the knee and hip extended. The current system only assesses plantar flexion in sitting so it does not capture the movement relevant to swimming where the ability to plantar flex the ankle is important for maintaining a streamline position and kicking (Hull, 1990).

The current study has also established preliminary normative values for a new ROM test battery designed for Para swimmers. Participants comprised athletes with health conditions including spinal cord injury, cerebral palsy and acquired brain injury, which can result in impaired ROM, impaired strength, impaired coordination, or a combination of these impairments. Athletes with disabilities had significantly less ROM than non-disabled participants in the majority of ROM tests (Table 4). These results indicate that divergent validity is evident within this test battery. That is, we would expect people with impaired ROM to have lower scores than non-disabled people. Further analysis indicated a difference within the group of athletes with disabilities. When comparing those with either predominately impaired ROM associated with hypertonia (resulting from cerebral palsy and acquired brain injury) and those with impaired muscle power (resulting primarily from spinal cord injury) there were differences in ROM at the trunk, hip and knee. These differences are not surprising given that all athletes with impaired muscle power had lesions that would limit their ability to actively move or control their trunk and lower limbs. This outcome implies some degree of discriminant ability within the test battery, but it should be noted that despite finding some differences in the pattern of ROM between subgroups of athletes, the purpose of Paralympic classification is to identify and measure impairment, rather than classify merely based on a medical diagnosis (Tweedy & Vanlandewijck, 2011). Further research is needed to assess how impairments in ROM impact swimming performance.

The final feature of this study was that the majority of assessments in the new test battery were measured in a group of 24 Para swimmers. The study identified that the new test procedures are feasible within a sample of athletes with disabilities – including those with severe impairments. For example, two participants with impaired muscle power who compete in class S1 – the class for those with the most severe and limiting impairments - were able to complete the majority of tests in the test battery. One of the major limitations of the study was that the Para swimmers did not complete the supported prone position assessments of shoulder ROM (shoulder extension, horizontal abduction and unilateral flexion) due to time constraints and limited personnel. These data were collected during training camps so, unfortunately, it was inevitable that some measures could not be taken. As such, it is not yet known whether these supported prone positions are easily administered in swimming athletes with a disability or whether these measures produce systematic differences in ROM between non-disabled participants and swimming athletes with a disability. Additionally, as can be seen with the large ICC 95% confidence intervals for some measures (Tables 2 and 3), the sample size for the reliability measures is smaller than ideal to provide meaningful minimum detectable changes (Lexell & Downham, 2005) although the sample used for the respective reliability measures (intra- and inter-examiner) were very similar to those that have previously assessed the reliability of shoulder ROM (Bak & Magnusson, 1997; Furness, et al., 2015; Walker, et al., 2016). A further limitation was that inclinometer measurements of knee flexion and elbow ROM were only included for inter-examiner testing. Additionally, as this test battery focused on the assessment of motions relevant to S-class swimming, it is likely that additional tests addressing hip ~~external~~ rotation, ankle dorsi-flexion and ~~forearm~~ ankle supination are required for accurate SB-class (breaststroke) classification. Finally, although these tests were found to be reliable in a non-disabled population, reliability is a population specific characteristic, therefore the reliability of these tests needs to be confirmed in a disabled population before their utility in the classification process is established.

Conclusion

Overall, the novel ROM tests assessed for this study had good to excellent intra-examiner and inter-examiner reliability in non-disabled participants. In addition to evaluating the reliability for these ROM measures, preliminary normative ROM values for both non-disabled participants and Para swimmers have also been established. Finally, this study identified that Para swimmers had significantly less ROM than non-disabled participants. Further research should focus on the reliability of these tests in Para swimmers and this test battery should be validated against swimming performance in both abled-bodied and Para swimmers to determine which tests or combination of tests best predict swimming performance.

References

Awan, R., Smith, J., & Boon, A. J. (2002). Measuring shoulder internal rotation range of motion: a comparison of 3 techniques. *Archives of Physical Medicine and Rehabilitation, 83*, 1229-1234.

Bak, K., & Magnusson, S. P. (1997). Shoulder strength and range of motion in symptomatic and pain-free elite swimmers. *American Journal of Sports Medicine, 25*, 454-459.

Beckman, E. M., Connick, M. J., & Tweedy, S. M. (2016). How much does lower body strength impact Paralympic running performance? *European Journal of Sport Science, 16*, 669-676.

Beckman, E. M., & Tweedy, S. M. (2009). Towards evidence-based classification in Paralympic athletics: evaluating the validity of activity limitation tests for use in classification of Paralympic running events. *British Journal of Sports Medicine, 43*, 1067-1072.

Blanch, P. (2004). Conservative management of shoulder pain in swimming. *Physical Therapy in Sport, 5*, 109-124.

Boon, A. J., & Smith, J. (2000). Manual scapular stabilization: its effect on shoulder rotational range of motion. *Archives of Physical Medicine and Rehabilitation, 81*, 978-983.

Chatard, J. C., Lavoie, J. M., Bourgoin, B., & Lacour, J. R. (1990). The contribution of passive drag as a determinant of swimming performance. *International Journal of Sports Medicine, 11*, 367-372.

Cools, A. M., De Wilde, L., Van Tongel, A., Ceyssens, C., Ryckewaert, R., & Cambier, D. C. (2014). Measuring shoulder external and internal rotation strength and range of motion: comprehensive intra-rater and inter-rater reliability study of several testing protocols. *Journal of Shoulder and Elbow Surgery, 23*, 1454-1461.

Currier, L. L., Froehlich, P. J., Carow, S. D., McAndrew, R. K., Cliborne, A. V., Boyles, R. E., Mansfield, L. T., & Wainner, R. S. (2007). Development of a clinical prediction rule to identify patients with knee pain and clinical evidence of knee osteoarthritis who demonstrate a favorable short-term response to hip mobilization. *Physical Therapy, 87*, 1106-1119.

Daly, D. J., Malone, L. A., Smith, D. J., Vanlandewijck, Y. C., & Steadward, R. D. (2001). The contribution of starting, turning, and finishing to total race performance in male Paralympic swimmers. *Adapted Physical Activity Quarterly, 18*, 316-333.

De Winter, A. F., Heemskerk, M., Terwee, C. B., Jans, M. P., Devillé, W., Van Schaardenburg, D., Scholten, R. J., & Bouter, L. M. (2004). Inter-observer reproducibility of measurements of range of motion in patients with shoulder pain using a digital inclinometer. *BMC Musculoskeletal Disorders, 5*, 18.

Dummer, G. M. (1999). Classification of swimmers with physical disabilities. *Adapted Physical Activity Quarterly, 16*, 216-218.

Elveru, R. A., Rothstein, J. M., & Lamb, R. L. (1988). Goniometric reliability in a clinical setting: subtalar and ankle joint measurements. *Physical Therapy, 68*, 672-677.

Evershed, J., Frazer, S., Mellifont, R. B., & Burkett, B. (2012). Sports technology provides an objective assessment of the Paralympic swimming classification system. *Sports Technology, 5*, 49-55.

Furness, J., Johnstone, S., Hing, W., Abbott, A., & Climstein, M. (2015). Assessment of shoulder active range of motion in prone versus supine: a reliability and concurrent validity study. *Physiotherapy Theory and Practice, 31*, 489-495.

Howe, P. D., & Jones, C. (2006). Classification of disabled athletes:(Dis) empowering the Paralympic practice community. *Sociology of Sport Journal, 23*, 29-46.
Hull, M. (1990). Flexible ankles: Faster swimming. *Swimming Technique, 27*, 23-24.
IPC. (2007). International Paralympic Committee (IPC) classification code and international standards. In (Vol. 2017): IPC.
IPC. (2016). Rio 2016 Paralympic ticket sales exceed 1.8 million. In: International Paralympic Committee
IPC. (2017). World Para Swimming Classification Rules and Regulations In. Bonn, Germany.
Kolber, M. J., Vega Jr, F., Widmayer, K., & Cheng, M. S. (2011). The reliability and minimal detectable change of shoulder mobility measurements using a digital inclinometer. *Physiotherapy Theory and Practice, 27*, 176-184.
Lexell, J. E., & Downham, D. Y. (2005). How to assess the reliability of measurements in rehabilitation. *American Journal of Physical Medicine & Rehabilitation, 84*, 719-723.
Muir, S. W., Corea, C. L., & Beaupre, L. (2010). Evaluating change in clinical status: reliability and measures of agreement for the assessment of glenohumeral range of motion. *North American Journal of Sports Physical Therapy, 5*, 98-110.
Oh, Y. T., Burkett, B., Osborough, C., Formosa, D., & Payton, C. (2013). London 2012 Paralympic swimming: passive drag and the classification system. *British Journal of Sports Medicine, 47*, 838-843.
Payton, C. J., Bartlett, R. M., Baltzopoulos, V., & Coombs, R. (1999). Upper extremity kinematics and body roll during preferred-side breathing and breath-holding front crawl swimming. *Journal of Sports Sciences, 17*, 689-696.
Pink, M., Perry, J., Browne, A., Scovazzo, M. L., & Kerrigan, J. (1991). The normal shoulder during freestyle swimming: an electromyographic and cinematographic analysis of twelve muscles. *American Journal of Sports Medicine, 19*, 569-576.
Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research: applications to practice.* (3rd ed.). New Jersey: Pearson Education Inc.
Riemann, B. L., Witt, J., & Davies, G. J. (2011). Glenohumeral joint rotation range of motion in competitive swimmers. *Journal of Sports Sciences, 29*, 1191-1199.
Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.
Tweedy, S. M., Beckman, E. M., & Connick, M. J. (2014). Paralympic classification: conceptual basis, current methods, and research update. *PM&R, 6*, S11-S17.
Tweedy, S. M., & Vanlandewijck, Y. C. (2011). International Paralympic Committee position stand—background and scientific principles of classification in Paralympic sport. *British Journal of Sports Medicine, 45*, 259-269.
van de Pol, R. J., van Trijffel, E., & Lucas, C. (2010). Inter-rater reliability for measurement of passive physiological range of motion of upper extremity joints is better if instruments are used: a systematic review. *Journal of Physiotherapy, 56*, 7-17.
van Trijffel, E., van de Pol, R. J., Oostendorp, R. A. B., & Lucas, C. (2010). Inter-rater reliability for measurement of passive physiological movements in lower extremity joints is generally low: a systematic review. *Journal of Physiotherapy, 56*, 223-235.
Walker, H., Pizzari, T., Wajswelner, H., Blanch, P., Schwab, L., Bennell, K., & Gabbe, B. (2016). The reliability of shoulder range of motion measures in competitive swimmers. *Physical Therapy in Sport, 21*, 26-30.

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of Strength and Conditioning Research, 19*, 231-240.

Youdas, J. W., Bogard, C. L., & Suman, V. J. (1993). Reliability of goniometric measurements and visual estimates of ankle joint active range of motion obtained in a clinical setting. *Archives of Physical Medicine and Rehabilitation, 74*, 1113-1118.

Zamparo, P., Gatta, G., Pendergast, D., & Capelli, C. (2009). Active and passive drag: the role of trunk incline. *European Journal of Applied Physiology, 106*, 195-205.

Table 1: Characteristics of non-disabled participants and Para swimmers

	Non-disabled	Swimmers with hypertonia	Swimmers with impaired muscle power
n	42	11	13
Health condition	N/A	CP (n=10), ABI (n=1)	SCI (n=12), polio (n=1),
Age (years)	23.2 (4.5)	27 (5.7)	30 (6.4)
Height (m)	1.74 (0.18)	1.66 (0.16)	1.62 (0.14)
Mass (kg)	73.5 (11.5)	66.6 (10.7)	60.5 (13.5)
BMI (kg/m^2)	23.4 (2.9)	24.3 (3.6)	22.8 (2.6)

Values presented as mean (SD); BMI = body mass index; N/A = not applicable; CP = cerebral palsy; ABI = acquired brain injury; SCI = spinal cord injury

Table 2: Intra-examiner reliability for ROM measures in non-disabled participants

Measure	ICC	ICC 95% CI	SEM	T test p value
Bilateral Shoulder flexion (I)	0.95	0.86-0.98	2.96	0.39
Bilateral Shoulder abduction (I)	0.73	0.24-0.91	3.85	0.22
Elbow flexion L (G)	0.82	0.48-0.94	1.69	0.19
Elbow flexion R (G)	0.81	0.46-0.94	1.56	0.40
Elbow extension L (G)	0.95	0.84-0.98	1.84	0.04
Elbow extension R (G)	0.94	0.81-0.98	1.42	0.95
Lower-limb streamline thigh L (I)	0.91	0.75-0.97	1.18	0.51
Lower-limb streamline thigh R (I)	0.83	0.51-0.94	1.80	0.25
Lower-limb streamline shank L (I)	0.89	0.69-0.96	1.31	0.63
Lower-limb streamline shank R (I)	0.91	0.71-0.97	1.26	0.87
Lower-limb streamline L knee (I)	0.94	0.82-0.98	1.40	0.31
Lower-limb streamline R knee (I)	0.91	0.75-0.97	1.61	0.26
Lower-limb streamline L ankle plantar-flexion (I)	0.95	0.57-0.99	1.67	$\leq$0.01
Lower-limb streamline R ankle plantar-flexion (I)	0.96	0.89-0.99	1.73	0.22
Hip flexion L (I)	0.95	0.85-0.98	2.91	0.65
Hip flexion R (I)	0.92	0.76-0.97	3.45	0.24
Knee flexion L (G)	0.95	0.81-0.98	1.33	0.04
Knee flexion R (G)	0.88	0.64-0.96	2.64	0.38
Shoulder internal rotation L (I)	0.91	0.72-0.97	3.48	0.83
Shoulder internal rotation R (I)	0.88	0.64-0.96	4.00	0.61
Shoulder external rotation L (I)	0.83	0.48-0.94	6.11	0.56
Shoulder external rotation R (I)	0.81	0.43-0.94	6.29	0.91
Prone shoulder horizontal abduction L (I)	0.87	0.62-0.96	3.07	0.56
Prone shoulder horizontal abduction R (I)	0.82	0.48-0.94	3.26	0.14
Prone shoulder extension, elbow flexed L (I)	0.94	0.82-0.98	3.09	0.70
Prone shoulder extension, elbow flexed R (I)	0.90	0.68-0.97	3.95	0.17
Prone shoulder extension, elbow extended L (I)	0.94	0.82-0.98	3.46	0.31
Prone shoulder extension, elbow extended R (I)	0.90	0.71-0.97	4.60	0.79
Prone shoulder flexion L (I)	0.96	0.87-0.99	2.37	0.15
Prone shoulder flexion R (I)	0.90	0.71-0.97	3.61	0.80
Functional reach Forwards (T)	0.66	-0.02-0.89	2.24	0.48
Functional reach L (T)	0.80	0.41-0.93	2.00	0.31
Functional reach R (T)	0.68	0.07-0.89	2.35	0.31

Functional reach backwards (T)	0.94	0.83-0.98	2.51	0.37

L = left; R = right; (I) = measurement with inclinometer; (G) = measurement with goniometer; (T) = measurement with tape measure; ICC = intraclass correlation coefficient; ICC 95% CI = intraclass correlation coefficient 95% confidence intervals; SEM = standard error of measurement; T test conducted on mean values from session one and two.

Table 3: Inter-examiner reliability for ROM measures in non-disabled participants

Measure	ICC	ICC 95% CI	SEM	T-test p value
Bilateral Shoulder flexion (I)	0.98	0.94-0.99	1.42	0.99
Bilateral Shoulder abduction (I)	0.85	0.59-0.95	2.74	0.20
Elbow flexion L (G)	0.79	0.17-0.94	2.04	$\leq$0.01
Elbow flexion R (G)	0.67	0.10-0.88	3.05	0.03
Elbow extension L (G)	0.76	0.30-0.92	1.88	0.03
Elbow extension R (G)	0.73	0.17-0.91	2.12	0.01
Elbow flexion L (I)	0.82	0.13-0.95	3.81	$\leq$0.01
Elbow flexion R (I)	0.86	0.61-0.95	3.79	0.18
Elbow extension L (I)	0.90	0.70-0.97	3.19	0.09
Elbow extension R (I)	0.75	0.26-0.91	3.94	0.93
Lower-limb streamline thigh L (I)	0.92	0.77-0.97	1.10	0.17
Lower-limb streamline thigh R (I)	0.92	0.77-0.97	0.73	0.55
Lower-limb streamline shank L (I)	0.89	0.67-0.96	1.44	0.76
Lower-limb streamline shank R (I)	0.80	0.42-0.93	1.77	0.50
Lower-limb streamline L knee (I)	0.88	0.68-0.96	0.88	0.31
Lower-limb streamline R knee (I)	0.79	0.41-0.93	0.79	0.41
Lower-limb streamline L ankle plantar-flexion (I)	0.98	0.94-0.99	0.98	0.65
Lower-limb streamline R ankle plantar-flexion (I)	0.98	0.95-0.99	0.98	0.84
Hip flexion L (I)	0.88	0.57-0.96	3.14	0.02
Hip flexion R (I)	0.92	0.76-0.97	2.40	0.70
Knee flexion L (G)	0.21	-0.71-0.69	3.50	0.04
Knee flexion R (G)	0.72	0.23-0.90	2.43	0.26
Knee flexion L (I)	0.85	0.55-0.95	3.09	0.97
Knee flexion R (I)	0.82	0.48-0.94	3.51	0.52
Shoulder internal rotation L (I)	0.86	0.42-0.96	6.52	0.01
Shoulder internal rotation R (I)	0.86	0.50-0.95	6.34	0.02
Shoulder external rotation L (I)	0.92	0.78-0.97	4.96	0.13
Shoulder external rotation R (I)	0.95	0.85-0.98	4.22	0.97
Supported internal rotation L (I)	0.81	0.45-0.94	5.73	0.95
Supported internal rotation R (I)	0.90	0.71-0.97	3.80	0.60
Prone shoulder horizontal abduction L (I)	0.93	0.79-0.97	3.85	0.91
Prone shoulder horizontal abduction R (I)	0.86	0.62-0.95	4.80	0.31
Prone shoulder extension, elbow flexed L (I)	0.96	0.89-0.99	2.75	0.93
Prone shoulder extension, elbow flexed R (I)	0.95	0.87-0.98	2.79	0.05
Prone shoulder extension, elbow extended L (I)	0.96	0.88-0.99	2.99	0.61
Prone shoulder extension, elbow extended R (I)	0.96	0.88-0.99	2.65	0.61
Prone shoulder flexion L (I)	0.84	0.55-0.94	3.36	0.40

Prone shoulder flexion R (I)	0.95	0.85-0.98	2.24	0.80
Functional reach forwards (T)	0.73	0.21-0.91	2.74	0.75
Functional reach L (T)	0.76	0.29-0.92	2.45	0.11
Functional reach R (T)	0.62	0.02-0.86	2.96	0.94
Functional reach backwards (T)	0.92	0.76-0.92	2.24	0.08

L = left; R = right; (I) = measurement with inclinometer; (G) = measurement with goniometer; (T) = measurement with tape measure; ICC = intraclass correlation coefficient; ICC 95% CI = intraclass correlation coefficient 95% confidence intervals; SEM = standard error of measurement; T test conducted on mean values from examiner one and two.

Table 4: Range of motion preliminary normative values for entire test battery for non-disabled participants and Para swimmers

Measure	**Non-disabled** ROM degrees (SD)	**Swimmers with Hypertonia** ROM degrees (SD)	**Swimmers with impaired muscle power** ROM degrees (SD)
Bilateral Shoulder flexion (I)	165 (12)	151 (15)	147 (31)
$^{\beta}$Bilateral Shoulder abduction (I)	184 (8)	186 (12)	179 (20)
Elbow flexion L (G)	143 (5)	135 (7)	132**(21)
Elbow flexion R (G)	144 (6)	136 (8)	128** (33)
Elbow extension L (G)	-1.4 (6)	-11** (12)	-8.4* (8)
Elbow extension R (G)	-2.3 (6)	-6.6 (8)	-8.0* (9)
Elbow flexion left (I)	146 (9)		
Elbow flexion right (I)	146 (11)		
Elbow extension left (I)	4.7 (9)		
Elbow extension right (I)	4.8 (11)		
Elbow total ROM left (I)	151 (17)		
Elbow total ROM right (I)	151 (15)		
$^{\beta}$Lower-limb streamline thigh L (I)	8.1 (7)	2.0 (10)	3.7 (17)
$^{\beta}$Lower-limb streamline thigh R (I)	6.7 (6)	0.9 (11)	2.0 (18)
$^{\beta}$Lower-limb streamline shank L (I)	5.3 (5)	-16** (13)	-15** (20)
$^{\beta}$Lower-limb streamline shank R (I)	5.4 (6)	-13** (11)	-14** (20)
$^{\beta}$Lower-limb streamline L knee (I)	2.8 (6)	18* (22)	19* (37)
$^{\beta}$Lower-limb streamline R knee (I)	1.3 (5)	14 (22)	15 (38)
$^{\beta}$Lower-limb streamline L ankle plantar-flexion (I)	9.5 (10)	40** (19)	33** (12)
$^{\beta}$Lower-limb streamline R ankle plantar-flexion (I)	7.9 (10)	38** (21)	33** (16)
Hip flexion L (I)	119 (13)	69** (32)	44** (54)
Hip flexion R (I)	118 (13)	67**^ (31)	36** (46)
Knee flexion L (G)	135 (7)	100**^ (34)	59 ** (63)
Knee flexion R (G)	134 (8)	101**^ (27)	54** (64)
Knee flexion L (I)	36 (8)		
Knee flexion R (I)	37 (8)		
Shoulder internal rotation L (I)	83 (17)	36**^ (19)	56** (13)
Shoulder internal rotation R (I)	78 (16)	44** (17)	55** (13)
Shoulder external rotation L (I)	92 (15)	82 (18)	77* (25)
Shoulder external rotation R (I)	97 (15)	88 (11)	85 (20)
$^{\beta}$Shoulder internal rotation supported L (I)	89 (11)		
$^{\beta}$Shoulder internal rotation supported R (I)	84 (10)		
$^{\beta}$Prone shoulder horizontal abduction L (I)	31 (12)		
$^{\beta}$Prone shoulder horizontal abduction R (I)	32 (12)		

$^{\beta}$Prone shoulder extension, elbow flexed L $^{(I)}$	44 (14)		
$^{\beta}$Prone shoulder extension, elbow flexed R $^{(I)}$	46 (14)		
Prone shoulder extension, elbow extended L $^{(I)}$	33 (16)		
Prone shoulder extension, elbow extended R $^{(I)}$	35 (14)		
$^{\beta}$Prone shoulder flexion L $^{(I)}$	169 (12)		
$^{\beta}$Prone shoulder flexion R $^{(I)}$	167 (12)		
$^{\beta}$Functional reach forwards $^{(T)}$	54 (7)	42*^ (11)	15** (20)
$^{\beta}$Functional reach L $^{(T)}$	23 (5)	15**^ (4)	7.9** (10)
$^{\beta}$Functional reach R $^{(T)}$	24 (5)	15**^ (3)	8.0** (9)
$^{\beta}$Functional reach backwards $^{(T)}$	35 (10)	17**^ (5)	6.9** (9)

Values presented as mean (SD) for degrees of range of motion or distance (cm) for functional reach measures. L = left; R = right; $^{(I)}$ = measurement with inclinometer; $^{(G)}$ = measurement with goniometer; $^{(T)}$ = measurement with tape measure; *significantly different to non-disabled participants ($p \leq 0.05$); ** significantly different to non-disabled participants ($p \leq 0.01$); ^ significantly different from athletes with impaired muscle power ($p \leq 0.05$). $^{\beta}$ designate novel tests that are not in the current classification system (all tests are described in supplementary table 1). Area in grey represents tests that were not completed by Para swimmers due to time/personnel restrictions.