Musculoskeletal Science and Practice 40 (2019) 45–52

Contents lists available at ScienceDirect

Musculoskeletal Science and Practice

journal homepage: www.elsevier.com/locate/msksp

Original article

Reliability and validity of a mobile tablet for assessing left/right judgements

Lindy J. Williams[a,*], Felicity A. Braithwaite[a], Hayley B. Leake[a], Michelle N. McDonnell[a], Daniela K. Peto[a], G. Lorimer Moseley[a,b], Susan L. Hillier[a]

[a] University of South Australia, Adelaide, South Australia, 5000, Australia
[b] Neuroscience Research Australia, Sydney, Australia

ARTICLE INFO

Keywords:
left/right judgement
Motor imagery
Body schema
Concurrent validity

ABSTRACT

Background: Left/right judgement (LRJ) of body parts is commonly used to assess the ability to perform implicit motor imagery and the integrity of brain-grounded maps of the body. Clinically, LRJ are often undertaken using a mobile tablet, but the concurrent validity and reliability of this approach has not yet been established.
Objectives: To evaluate the concurrent validity and test-retest reliability of a mobile tablet for assessing LRJ.
Method: Participants completed LRJ for 50 hand images (Experiment 1), and 40 back, foot, or neck images (Experiment 2) using a mobile tablet and desktop computer in random order. Participants in Experiment 2 performed a repeat test the following day to assess test-retest reliability. Accuracy and response time (RT) were recorded.
Results: Twenty participants aged 55.3 (± 6.7) years in Experiment 1, and 37 participants aged 38.2 (± 12.3) years in Experiment 2, were recruited. Concurrent validity of the mobile tablet was good to excellent for hand judgements ($ICC_{3,1}$ = 0.836 for RT; ICC = 0.909 for accuracy), and was good for back, foot, and neck judgements (ICC = 0.781 for accuracy; ICC = 0.880 for RT). Test-retest reliability of the mobile tablet was good to excellent (ICC = 0.824 for accuracy; ICC = 0.903 for RT).
Conclusions: The mobile tablet demonstrated good to excellent concurrent validity with the desktop computer in two separate samples. The mobile tablet also demonstrated good to excellent test-retest reliability. The mobile tablet for LRJ is a valid alternative to the original desktop version.

1. Introduction

Left/right judgements (LRJ) (Parsons, 1987) require a judgement as to whether a pictured body part belongs to, or has moved towards, the left or right. The process is thought to involve three steps: firstly an unconscious decision of left or right is made, then the chosen side is mentally rotated into the same alignment for comparison, and finally the initial choice is accepted or rejected (Parsons, 1987). Whether or not this process also applies to neck and trunk images has not been determined. The outputs of the task are response time (RT) and accuracy. As the task occurs without physically moving the tested body part, it is frequently used to assess motor imagery (Botnmark et al., 2016; Breckenridge et al., 2017; Wallwork et al., 2013), which relies on the integrity of brain-grounded maps – or cortical representations – of the body, receiving inputs from various brain areas that process visual, proprioceptive and tactile input (Hétu et al., 2013).

Recent evidence suggests that cortical representations of the body are altered in many different conditions, as assessed by LRJ performance, including complex regional pain syndrome (CRPS) (Moseley, 2004), phantom limb pain (Nico et al., 2004), knee osteoarthritis (Stanton et al., 2012), back pain (Bowering et al., 2014; Bray and Moseley, 2011), chronic facial pain (von Piekartz et al., 2014), chronic arm pain (Coslett et al., 2010b), chronic leg pain (Coslett et al., 2010a), cervical dystonia (Bradnam et al., 2015; Fiorio et al., 2007), carpal tunnel syndrome (Schmid and Coppieters, 2012), stroke (Amesz et al., 2016) and neglect (Baas et al., 2011; Coslett, 1998; Reinhart et al., 2012). As such, therapies aimed at normalising LRJ, such as Graded Motor Imagery (Moseley et al., 2012), have evolved and show promising results at reducing pain and disability (Bowering et al., 2013) (although see (Johnson et al., 2012) for no effect as part of a wider multimodal program). LRJ could potentially be used to facilitate motor recovery following stroke (Page et al., 2007) as a person with hemiplegia is still able to use motor imagery to mentally simulate movements of their affected limb (Johnson et al., 2002).

The Recognise™ (Neuro Orthopaedic Institute and Noigroup Publications, 2012) software allows for quick and easy assessment of

* Corresponding author. School of Health Sciences, University of South Australia, GPO Box 2471, Adelaide, SA, 5001, Australia.
E-mail address: lindy.williams@unisa.edu.au (L.J. Williams).

https://doi.org/10.1016/j.msksp.2019.01.010
Received 8 August 2018; Received in revised form 15 January 2019; Accepted 19 January 2019

LRJ. Images displayed on a screen show a left or right limb – foot, hand, shoulder or knee – or an axial image – back or neck - rotated or tilted to the left or right. The number of total images presented can be customised from 6, 10, 20, 30, 40 or 50. The images are presented in random orientations (0°, 90°, 180°, 270°), and total accuracy and average RT are calculated and presented at the end of the task.

The reliability of the desktop computer version of Recognise™ has previously been established in people with and without pain. The accuracy score has good test-retest reliability with 40 hand *(ICC = 0.924; 0.89)* and back images *(ICC = 0.92; 0.80),* with moderate to good reliability of RT score of hand *(ICC = 0.697; 0.83)* and Back LRJ *(ICC = 0.87; 0.74)* (Bray and Moseley, 2011; Dey et al., 2012).

Substantial limitations of the desktop computer version, such as limited portability and the requirement of access to the internet or Wi-Fi have been overcome by the mobile tablet version of Recognise™ designed for tablets and smartphones. The mobile tablet version of Recognise™ presents a change in the hardware being used, however there are no programmed differences with the software. The device is often held by one hand or placed horizontally in front of the user; the images are usually smaller; the responses involve touch screen interface instead of keyboard.

The mobile tablet version of Recognise is being used clinically and within research (Wong and Wong, 2017; Zimney et al., 2018), but the reliability and validity of this method has not yet been established. No studies have investigated the reliability or validity of tablet based back, neck and foot LRJ. One study investigating the reliability of hand LRJ in comparison to card-based measurement suggests good reliability (ICC = 0.897) with RT, however, the accuracy score was poor to moderate (ICC = 0.382; 0.602) and varied between the different sessions (Zimney et al., 2018). This study did not use a practice test prior to the recorded tests, therefore the poor reliability for the accuracy score may have been due to participants being unfamiliar with the requirements of the test or a possible learning effect. Additionally, the participants in this study were younger adults with a narrow age range (24.3 years ± 2.5). Therefore, the results cannot be applied to an older population as evidence shows that accuracy (Dey et al., 2012) and RT (Saimpont et al., 2009) decrease as individuals age.

The aims of this study are to (1) evaluate the concurrent validity and test-retest reliability of a mobile tablet device for assessing LRJ in comparison to the original desktop version in a healthy adult population and (2) establish minimal detectable difference (MDD) for accuracy and reaction time for both methods using the Recognise iOS Apps (Recognise Hand, Recognise Back, Recognise Foot, and Recognise Neck).

2. Materials and methods

Written informed consent was given by all participants in Experiments 1 and 2, in accordance with the latest version of the Declaration of Helsinki, and ethical approval was obtained from the institutional ethics committee. Both experiments in this study have employed slightly different methods to explore the best protocol for LRJ measurement, as recommended by previous evidence (Zimney et al., 2018).

2.1. Experiment 1: hand judgements

Convenience sampling was used to recruit 20 participants over 40 years of age with no neurological or psychological diagnosis. The testing took place in a quiet office within a university. Random allocation to two groups using sealed envelope (Sealed Envelope Ltd, 2018), a web-based randomisation program, determined the order of assessments. Group *'A'* were assessed with the desktop version first, Group *'B'* were assessed with the mobile device first, with a break of approximately 1 min between the two tasks. For the desktop version, the keyboard was aligned on the desk directly in front of the seated participants at a comfortable position, with their face 50 cm from the screen. Participants were asked to place their hands palm down on the desk and use their "dominant hand to press the left arrow key on the keyboard when the image showed a left hand, and the right arrow when it showed a right hand". For the tablet version, the tablet was held by the assessor at an angle of 45° on the desk and approximately 50 cm from the participant's face. The screen displayed two buttons marked with left and right and the participant was asked to respond to each image by touching the corresponding button on the screen with their dominant hand.

Both versions displayed 50 images of right or left hands, on a plain mono-coloured background, in random orientations (0°, 90°, 180°, 270°) and postures. Responses automatically triggered the next image. If a response was not registered within 30 s, the next image was displayed. Each version displayed five pictures as a familiarisation process at the beginning, consistent with clinical practice, which did not count towards the results. The participants were blinded to the results of each trial including the familiarisation process until testing was complete. Accuracy score (correct responses as a percentage of number of trials) and average RT (seconds) were determined for each participant.

2.2. Experiment 2: back, neck and foot judgements

Convenience sampling was used to recruit 37 participants of any age. Sample size was determined by the available time and resources. A post-hoc analysis of sample size, revealed a minimum of 24 was satisfactory for a two-rater assessment, with width of the Wald interval at 0.15 assuming ratio of variability at 0.8 and providing a 50% assurance probability (Zou, 2012). Testing took place during intervals at a pain conference for health professionals. The body part to be assessed (back, neck or foot), and the order of assessments (desktop first or tablet first) were determined by blindly drawing from a hat to select device (desktop/tablet) and to select body part (back, neck or foot). To assess test-retest reliability, each participant completed the same test on the tablet again the next day.

The desktop and tablet were positioned on the table 50cm in front of the participant. Participants were asked to "press the left and right arrow keys on the keyboard" and to "touch the left or right button on the tablet screen" when they identified a left or right hand in the picture. Participants were advised that the Recognise™ App automatically calculates speed and accuracy of all responses. Each participant was fitted with headphones with white noise to minimise distraction due to auditory stimuli.

Both versions displayed 40 images of backs, necks or feet, on a plain mono-coloured background, in random orientations and positions, but with a counterbalanced mix of left and right. Each picture was shown for a maximum of 5 s. A familiarisation process (40 images) of the selected body part was completed immediately prior to commencing the recorded test for each device on the first day. The familiarisation process ensured that participants were familiar with the test and to account for any learning effect. This process was not completed for the repeat test due to the carry-over effect from the previous day. The participants were not blinded to the results of each test. Accuracy (correct responses as a percentage of number of trials) and average RT (seconds) were determined for each participant.

2.3. Statistical analysis

Full details of the *a priori* statistical analysis plan for both experiments are available on Open Science Framework (URL: https://osf.io/695j2). Any deviations from that plan are clearly marked here. Data were analysed using SPSS software (version 22 IBM Corp, Armonk, NY, USA). Baseline characteristics were summarised using descriptive statistics. Means, standard deviation (SD) and the 95% confidence intervals for assessing differences between the mobile device and the desktop version were calculated for each parameter.

The differences between assessments were checked for outliers (defined as observations that were greater than 3SD from the group mean difference) (Portney and Watkins, 2015, p.551), and the distributions of the differences between assessments were checked for normality and skewness (Bland and Altman, 1986). For Experiment 2, no significant differences were found with the regression coefficients of the three data sets (back, neck and foot LRJ accuracy and RT), therefore they were combined for the primary analyses.

To evaluate the concurrent validity of the mobile device in comparison to the desktop for assessing LRJ, all data were plotted against the line of equality in Microsoft Excel to graphically observe the differences between the mobile device and the desktop scores. For each primary outcome - Accuracy and RT - a two-way mixed model analysis was performed, and a single measure intraclass correlation coefficient ($ICC_{3,1}$) was computed to estimate the magnitude of the correlation (absolute agreement) between the two methods. Bland-Altman plots were constructed to check for systematic bias, variability and agreement. The agreement between the two methods was summarised using the mean difference and standard deviation of the difference, and the 95% limits of agreement were calculated (Bland and Altman, 1986).

To evaluate the test-retest reliability of the mobile device for assessing LRJ, back, neck and foot judgements, all data was plotted against the line of equality in Microsoft Excel to graphically observe the differences between the scores. For accuracy and RT, a two-way mixed model was performed, and a single measure intraclass correlation coefficient ($ICC_{3,1}$) was computed to estimate the magnitude of the reliability of the mobile device. Bland-Altman plots were constructed to check for systematic bias, variability, and agreement.

ICC was classified according to the Portney and Watkins' guidelines of excellent (> 0.90), good (0.75–0.90), moderate (0.5–0.75) and poor (< 0.50) (Portney and Watkins, 2015). To check for systematic bias, variability and agreement, Bland-Altman plots were constructed. The agreement between the two methods was summarised using the mean difference ($\bar{d}$) and standard deviation of the difference (s), and the 95% limits of agreement was calculated (Bland and Altman, 1986). The standard error of measurement (SEM) and minimal detectable difference (MDD) were calculated according to SEM = SD x (1- ICC), and MDD = SEM x 1.96 x √2 (Weir, 2005). The results were considered statistically significant with $p < 0.05$ *a priori*.

3. Results

Table 1 presents the baseline characteristics of participants in Experiments 1 and 2. Participants in Experiment 2 were asked if they had any persistent pain, out of whom 24% (9 out of 37) reported they did have persistent pain, which is representative of the wider population (Blyth et al., 2001; Fayaz et al., 2016; National Center for Health Statistics, 2016). All data were normally distributed. Two outliers were removed from the Experiment 2 data [one from the accuracy analysis comparing both methods (concurrent validity), and one from the RT analysis for Day 1 and 2 (test-retest reliability)], because the individual differences in scores were greater than three standard deviations from the mean difference.

Table 1
Baseline characteristics.

LRJ body part	Experiment 1 (n = 20)	Experiment 2 (n = 37)		
	Hand	Back (n = 12)	Neck (n = 11)	Foot (n = 14)
Age years), mean (SD)	55.3 (6.69)	41.67 (11.4)	40 (15)	33.79 (9.81)
Sex	15 Female 5 Male	8 Female 4 Male	9 Female 2 Male	11 F 3 M
Handedness	20 Right 0 Left	12 Right 0 Left	9 Right 2 Left	12 R 2 L

SD Standard Deviation.

Table 2
Concurrent validity for LRJ Accuracy and RT.

LRJ Task	Measure	n	F	ICC	ICC 95%CI	SEM	MDD
Hand	Accuracy (%)	20	12.147	0.909*	0.765–0.964	0.664	1.841
Hand	RT (sec)	20	6.060	0.836*	0.591–0.935	0.147	0.409
Back, Neck & Foot	Accuracy (%)	36	4.522	0.781*	0.571–0.888	1.006	2.789
Back, Neck & Foot	RT (sec)	37	8.134	0.880*	0.766–0.938	0.027	0.074

* = P < 0.0001.

3.1. Concurrent validity

The ICCs for accuracy and RT of hand LRJ were excellent and good respectively (see Table 2). Fig. 1 displays the Bland-Altman for the difference in accuracy and RT between the tablet and desktop LRJ for hands (upper panels) and back, neck and foot (lower panels). Linear regression analysis revealed no proportional bias for accuracy ($p = 0.590$) or RT ($p = 0.396$). Fig. 2 displays the line of equality for the two methods. The ICCs for accuracy and RT of back, neck and foot LRJ were good (Table 2) and no proportional biases were found from linear regression ($p = 0.199$ and $p = 0.625$, respectively).

3.2. Test-retest reliability

For test-retest reliability of the tablet device, the ICCs for accuracy and RT were good to excellent (Koo and Li, 2016) (Table 3). Linear regression revealed no proportional bias for RT ($p = 0.818$), but there was a proportional bias for accuracy ($p = 0.002$) with the back, neck and foot LRJ (see Fig. 3A). The Bland-Altman plots for the difference in accuracy and RT on the tablet between day one and two for back, neck and foot LRJ are displayed in Fig. 4. The Bland-Altman plot for accuracy (Fig. 4a) shows the proportional bias may exist for participants with less than 85% accuracy. This could indicate a trend may exist with participants performing proportionally worse on repeat LRJ's, if their initial accuracy score was less than 85%. Further inspection revealed all the data points below 85% accuracy were LRJ's of neck images.

The data for each body part were separated from the analysis to further explore the proportional bias in the back, neck and foot LRJ. This exploration was not part of our planned analysis (URL: https://osf.io/695j2). The linear regression analysis of the back and foot LRJ data revealed no proportional bias ($p = 0.647$). One sample *t*-test of the neck LRJ data for day 1 and day 2 revealed a significant difference ($p = 0.026$). Therefore, the regression analysis of neck LRJ could not be completed as the assumptions were not met. However, further regression analysis separating the data for left and right accuracy of neck LRJ revealed no proportional bias ($p = 0.755$; $p = 0.222$ respectively). This could indicate that the proportional bias does not exist and resulted from averaging the left and right judgements together.

4. Discussion

The aim of this study was to evaluate the concurrent validity and test-retest reliability for hand, back, neck and foot LRJ on a tablet device in comparison to the desktop version in healthy participants.

4.1. Concurrent validity

We found good to excellent ICC values when comparing the two methods for accuracy scores and RT, indicating high concurrent validity of the tablet device for assessing LRJ of the hands, as well as the back, neck and foot. This indicates that the tablet device version of the LRJ

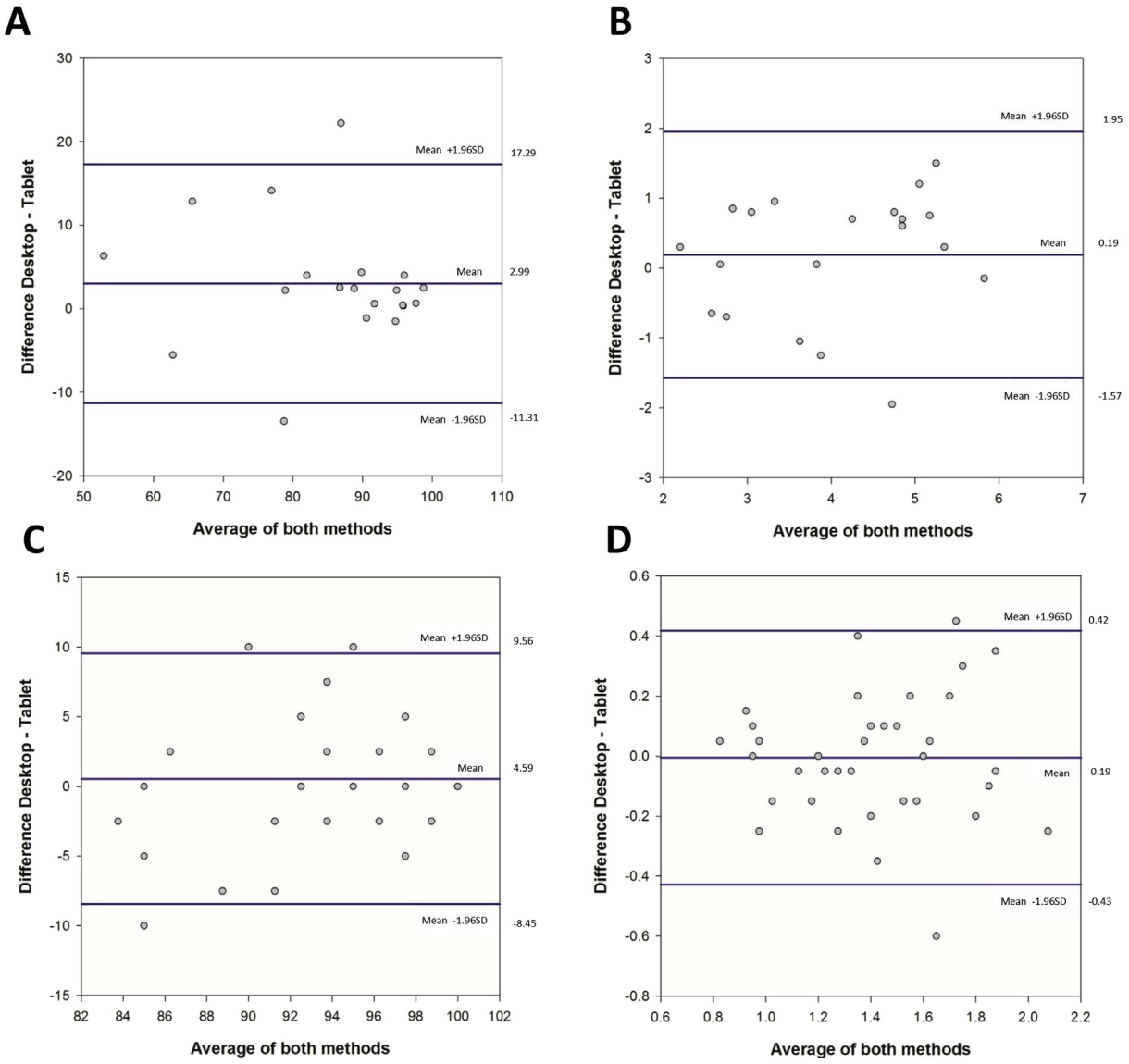

Fig. 1. Concurrent Validity. A, Bland-Altman plot for hand LRJ accuracy. B, Bland-Altman plot for hand LRJ reaction time. C, Bland-Altman plot for back, neck and foot LRJ accuracy. D, Bland-Altman plot for back, neck and foot LRJ reaction time.

task is a valid method for assessing LRJ of the hands, back, neck and foot.

4.2. Test retest reliability

The ICC values suggest the tablet device for assessing LRJ has good to excellent reliability for accuracy and RT scores. A proportional bias was found for test-retest reliability of the accuracy score for neck LRJ. This finding indicates that the two methods of assessing LRJ do not agree equally throughout the range of possible scores. For participants scoring less than 85% a possible trend may exist with lower accuracy resulting in proportionally worse scores on repeat LRJ's. This proportional bias cannot be explained by a learning effect (see (Boonstra et al., 2012)) as may be observed if the scores proportionally increased on repeat testing. We cannot be certain that the proportional bias exists. It may have been an artefact of averaging the left and right judgements together – as the bias was not evident when the left and right judgements were analysed separately.

If the bias does exist, a feasible explanation may relate to previous findings that healthy participants have greater variability and lower accuracy with left neck rotation judgements than with right neck rotation judgements (Leake, 2012; Wallwork et al., 2013). We have previously speculated (Wallwork et al., 2013) that this side to side difference may reflect a right neck rotation preference found in humans as reported by others (Güntürkün, 2003; Konishi et al., 1987; Ververs et al., 1994). Another possible explanation is that neck movements are highly integrated with vestibular data such that neck LRJ may be affected in a different way because of this. Hence neck LRJ may be a more complex task resulting in higher variability of accuracy scores with repeat testing, particularly with those participants that have a lower average accuracy score.

It is unlikely that the familiarisation process on day one influenced the results as a proportional bias was not found for the other body parts. The familiarisation process was to familiarise the participants with the test, and was therefore not required on repeat testing because observational data showed no performance loss over this time (GL Moseley Unpublished). However, the bias only existed for participants with low accuracy scores. Perhaps it cannot be assumed that participants with less than 85% accuracy for neck LRJ have a decent carry-over effect particularly considering that neck LRJ may be a more complex task, resulting in higher variability of accuracy scores with repeat testing (Leake, 2012; Wallwork et al., 2013). Therefore, the

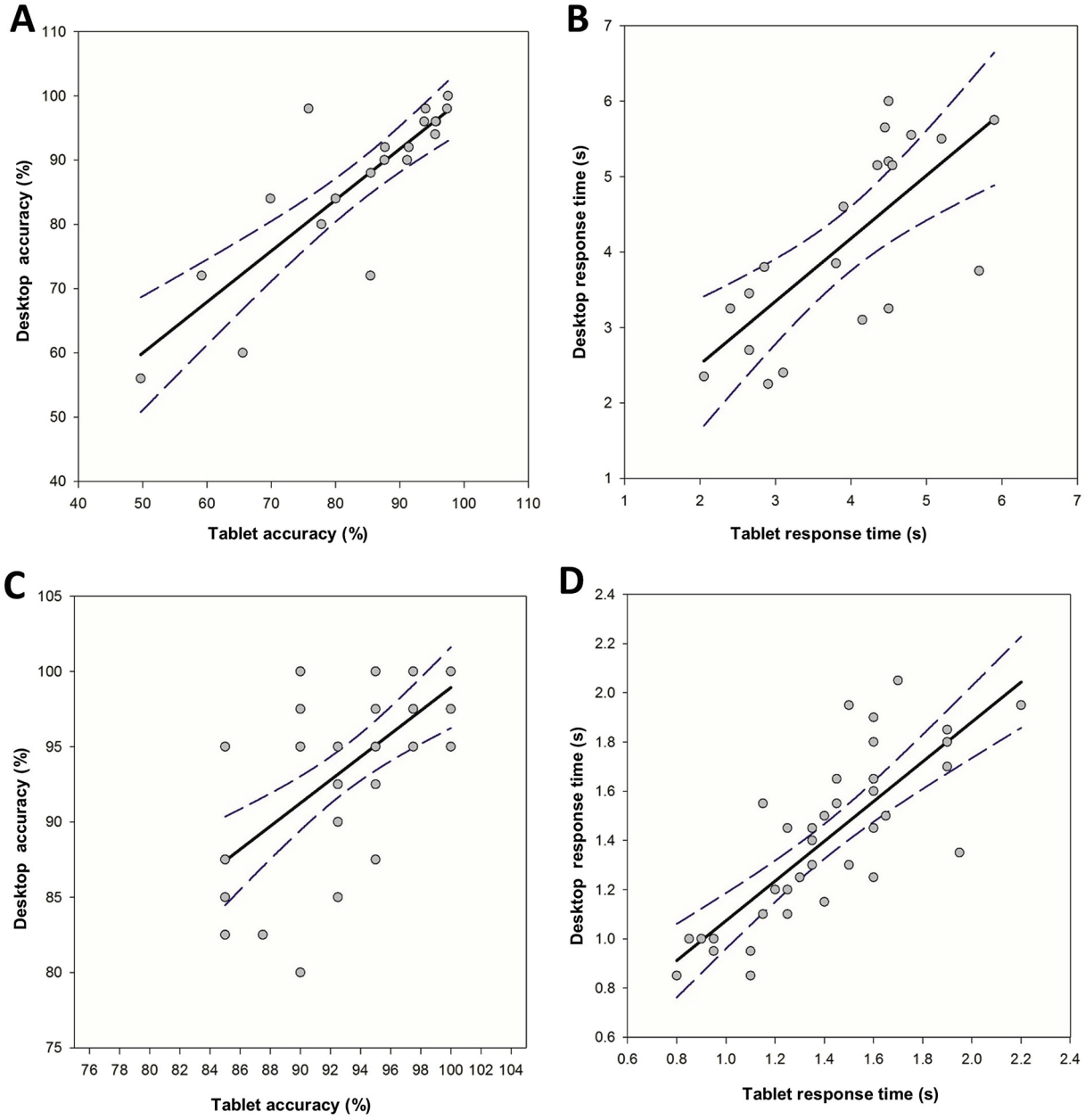

Fig. 2. Concurrent validity tablet versus desktop. A, Hand LRJ Accuracy. B, Hand LRJ response time. C, Back, neck & foot LRJ accuracy. D, Back, neck & foot LRJ response time.

Table 3
Test retest reliability for LRJ Accuracy and Response Time.

LRJ Task	Measure	n	F	ICC	ICC 95%CI	SEM	MDD
Back, Neck & Foot	Accuracy (%)	37	5.788	0.824[a]	0.661–0.909	0.846	2.345
Back, Neck & Foot	RT (sec)	36	12.386	0.903[a]	0.776–0.955	0.017	0.047

[a] = P < 0.0001.

familiarisation process may have influenced their scores.

Perhaps LRJ of the neck and back do not involve the same motor imagery process as limb LRJ. A recent report from a small study in healthy participants doing LRJ of back images, showed a negative relationship between mental movement amplitude and RT (Alazmi et al., 2018), which is opposite to that observed for limb LRJ. We did not detect this relationship in a previous study of over 1000 participants performing LRJ for trunk images (Bowering et al., 2014), nor in a study of over 1300 participants performing LRJ for neck images (Wallwork et al., 2013). Notably however, the Alazmi et al. (2018) study failed to provide practice trials, used both first and third person images and was underpowered, which leaves the matter unresolved. A proportionality bias has only been evident in previous back image data (see (Bowering et al., 2014)) (not neck image data, which is shown in this study) which suggests against the thesis that a different processing strategy for the task would explain the proportionality bias, but we certainly cannot unequivocally exclude the possibility.

Our assessment did not control for potential impact of images from different perspectives (e.g. (Meng et al., 2017) (note that some studies only include first person perspective images (e.g. (Bowering et al., 2014)), and we did not seek to evaluate any spatial effects of where on the screen the image occurred (see (Reid et al., 2016)). Despite this, the tests showed good to excellent agreement which lends confidence to our conclusion that they are valid and reliable.

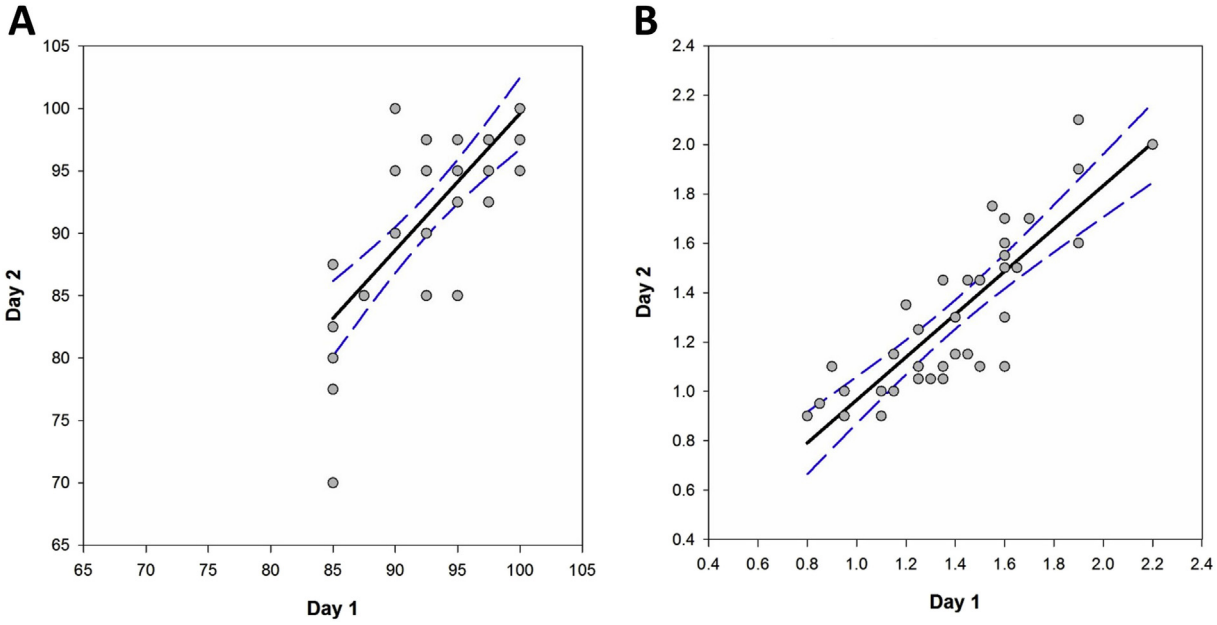

Fig. 3. A, Accuracy test-retest reliability: Day 1 versus Day 2. B, Response time test-retest reliability: Day 1 versus Day 2.

This study provides valuable information on the minimal detectable difference of tablet based LRJ's, which is critical for distinguishing real changes in repeat measurements from any random variation in the measurement tool itself (Haley and Fragala-Pinkham, 2006). Our findings have implications for clinical practice and research, particularly considering anecdotal evidence that tablets are widely used clinically for this purpose.

4.3. Strengths and limitations

We lodged and locked our *a priori* analysis plan prior to analysis, and only deviated from the plan to explore the proportionality bias, which is a strength, but we did not lock a full protocol prior to data collection, which is gold standard practice and now recommended for observational and behavioural studies (Lee et al., 2018). Both experiments involved reasonably small and convenience samples, although the slim confidence intervals around ICC's, means and standard deviations suggests that a larger sample would not have yielded a different response. Females were over represented in both our samples, which leaves open the question that the results may not be as reliable in males – we did not have enough males to undertake a sensitivity analysis on the basis of gender.

We pooled different body sites and acknowledge this may not be ideal. Future studies could aim for more participants per body site to test for differences. We found back LRJ had the highest average accuracy at 95%, neck was lowest at 89.5% and foot accuracy was in between at 91.4%. With small numbers we cannot be sure if the difference is significant or not.

We intentionally took slightly different approaches to the two experiments so as to evaluate several body parts and to more effectively

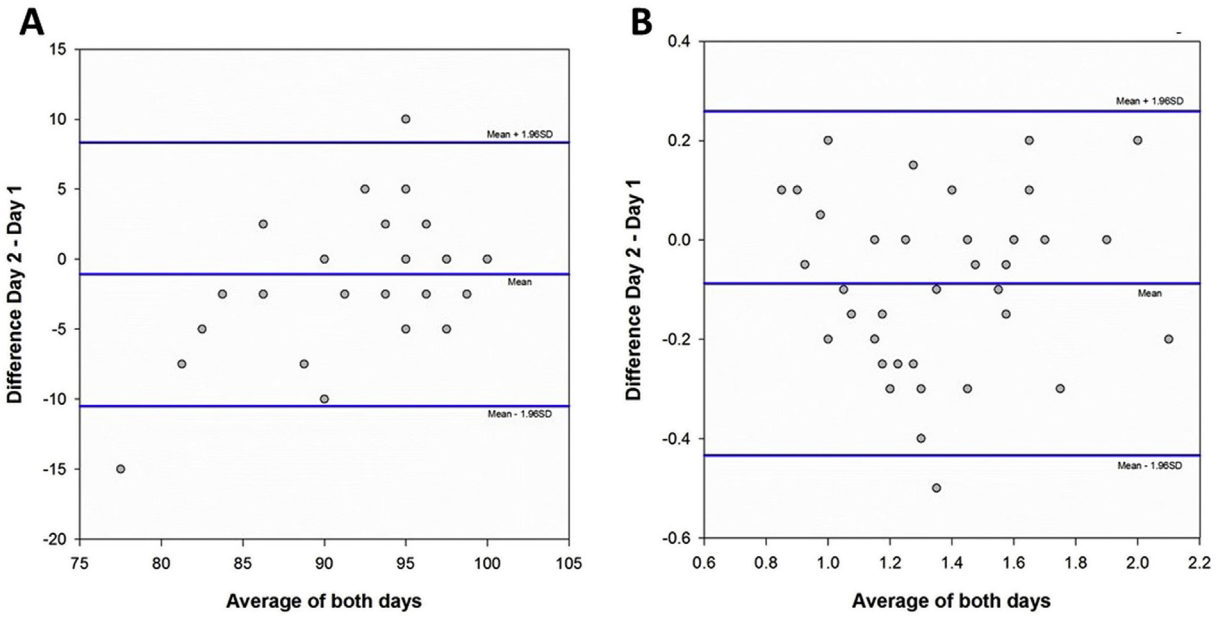

Fig. 4. Experiment 2 Test-retest reliability. A, Bland-Altman plot for back, neck and foot LRJ accuracy. B, Bland-Altman plot for back, neck and foot LRJ reaction time.

cover what we think is happening clinically and in research, but this prevented us from pooling data or looking specifically at sub-groups, which may have been useful considering the different age groups in the two experiments. Further related to protocol, Experiment 1 allowed participants five images for practice, while Experiment 2 allowed 40 images. It is our understanding that both approaches are employed clinically, although 40 images practice is recommended for the task to become implicit (Moseley, 2006) and we may have seen better validity and reliability for hand images if we had taken this approach. Finally, it should be remembered that we have established validity and reliability in healthy participants, some of whom had persistent pain (Experiment 2), not specific clinical populations.

5. Conclusions

The tablet version of the Recognise hand, back and foot LRJ tasks are valid and reliable and we can recommend their use in research and clinical practice in adult populations. The LRJ tasks may be less valid and reliable without a practise test of 40 images. We recommend caution when interpreting accuracy scores with neck LRJ, particularly with participants scoring lower than 85% accuracy. Further research is required to determine the validity and reliability in specific clinical populations.

Conflicts of interest

GLM has received support from Pfizer, Kaiser Permanente, Providence Health Care, Port Adelaide Football Club, Arsenal Football Club, the International Olympic Committee. He receives speaker fees for lectures on pain and rehabilitation and royalties for books on pain and rehabilitation, including one book directly relevant to this study – the Graded Motor Imagery Handbook. GLM receives fellowship support form NHMRC of Australia. HL was employed by Neuro Orthopaedic Institute during data collection but was not involved in analysis.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgements

Theo Niyonsenga for statistical advice for the development of the analysis protocol.

References

Alazmi, L., Gadsby, G.E., Heneghan, N.R., Punt, T.D., 2018. Do trunk-based left/right judgment tasks elicit motor imagery? Musculoskeletal Sci. Pract. 35, 55–60. https://doi.org/10.1016/j.msksp.2018.03.002.

Amesz, S., Tessari, A., Ottoboni, G., Marsden, J., 2016. An observational study of implicit motor imagery using laterality recognition of the hand after stroke. Brain Inj. 30 (8), 999–1004. https://doi.org/10.3109/02699052.2016.1147600.

Baas, U., de Haan, B., Grässli, T., Karnath, H.-O., Mueri, R., Perrig, W.J., Wurtz, P., Gutbrod, K., 2011. Personal neglect—A disorder of body representation? Neuropsychologia 49 (5), 898–905. https://doi.org/10.1016/j.neuropsychologia.2011.01.043.

Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1 (8476), 307–310 Feb 8.

Blyth, F.M., March, L.M., Brnabic, A.J.M., Jorm, L.R., Williamson, M., Cousins, M.J., 2001. Chronic pain in Australia: a prevalence study. Pain 89 (2), 127–134. https://doi.org/10.1016/S0304-3959(00)00355-9.

Boonstra, AMa, de Vries, SJbd, Veenstra, Ec, Tepper, Mc, Feenstra, Wac, Otten, Eb, 2012. Using the Hand Laterality Judgement Task to assess motor imagery: a study of practice effects in repeated measurements. Int. J. Rehabil. Res. 35 (3), 278–280. https://doi.org/10.1097/MRR.0b013e328355dd1e.

Botnmark, I., Tumilty, S., Mani, R., 2016. Tactile acuity, body schema integrity and physical performance of the shoulder: A cross-sectional study. Man. Ther. 23, 9–16. https://doi.org/10.1016/j.math.2016.02.001.

Bowering, K.J., Butler, D.S., Fulton, I.J., Moseley, G.L., 2014. Motor imagery in people with a history of back pain, current back pain, both, or neither. Clin. J. Pain 30 (12), 1070–1075. https://doi.org/10.1097/AJP.0000000000000066.

Bowering, K.J., O'Connell, N., Tabor, A., Catley, M., Leake, H., Moseley, G., Stanton, T., 2013. The effects of graded motor imagery and its components on chronic pain: a systematic review and meta-analysis. J. Pain 14 (1), 3–13. https://doi.org/10.1016/j.jpain.2012.09.007.

Bradnam, L.V., McDonnell, M.N., Ridding, M.C., 2015. Modulation of the cerebellum using theta-burst stimulation in people with cervical dystonia. Brain Stimul.: Basic Translat. Clin. Res. Neuromod. 8 (2), 329–330. https://doi.org/10.1016/j.brs.2015.01.068.

Bray, H., Moseley, G.L., 2011. Disrupted working body schema of the trunk in people with back pain. Br. J. Sports Med. 45 (3), 168–173. https://doi.org/10.1136/bjsm.2009.061978.

Breckenridge, J.D., McAuley, J.H., Butler, D.S., Stewart, H., Moseley, G.L., Ginn, K.A., 2017. The development of a shoulder specific left/right judgement task: Validity & reliability. Musculoskeletal Sci. Pract. 28, 39–45. https://doi.org/10.1016/j.msksp.2017.01.009.

Coslett, 1998. Evidence for a disturbance of the body schema in neglect. Brain Cogn. 37 (3), 527–544. https://doi.org/10.1006/brcg.1998.1011.

Coslett, H.B., Medina, J., Kliot, D., Burkey, A., 2010a. Mental motor imagery and chronic pain: The foot laterality task. J. Int. Neuropsychol. Soc. 16 (4), 603–612. https://doi.org/10.1017/S1355617710000299.

Coslett, H.B., Medina, J., Kliot, D., Burkey, A.R., 2010b. Mental motor imagery indexes pain: the hand laterality task. Eur. J. Pain 14 (10), 1007–1013.

Dey, A., Barnsley, N., Mohan, R., McCormick, M., McAuley, J.H., Moseley, G.L., 2012. Are children who play a sport or a musical instrument better at motor imagery than children who do not? Br. J. Sports Med. https://doi.org/10.1136/bjsports-2011-090525.

Fayaz, A., Croft, P., Langford, R.M., Donaldson, L.J., Jones, G.T., 2016. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open 6 (6). https://doi.org/10.1136/bmjopen-2015-010364.

Fiorio, M., Tinazzi, M., Ionta, S., Fiaschi, A., Moretto, G., Edwards, M., Bhatia, K., Aglioti, S., 2007. Mental rotation of body parts and non-corporeal objects in patients with idiopathic cervical dystonia. Neuropsychologia 45 (10), 2346–2354.

Güntürkün, O., 2003. Adult persistence of head-turning asymmetry. Nature 421, 711. https://doi.org/10.1038/421711a.

Haley, S.M., Fragala-Pinkham, M.A., 2006. Interpreting Change Scores of Tests and Measures Used in Physical Therapy. Phys. Ther. 86 (5), 735–743. https://doi.org/10.1093/ptj/86.5.735.

Hétu, S., Grégoire, M., Saimpont, A., Coll, M.-P., Eugène, F., Michon, P.-E., Jackson, P.L., 2013. The neural network of motor imagery: An ALE meta-analysis. Neurosci. Biobehav. Rev. 37 (5), 930–949. https://doi.org/10.1016/j.neubiorev.2013.03.017.

Johnson, S., Hall, J., Barnett, S., Draper, M., Derbyshire, G., Haynes, L., Rooney, C., Cameron, H., Moseley, G.L., de, C.W.A.C., McCabe, C., Goebel, A., 2012. Using graded motor imagery for complex regional pain syndrome in clinical practice: failure to improve pain. Eur. J. Pain 16 (4), 550–561. https://doi.org/10.1002/j.1532-2149.2011.00064.x.

Johnson, S.H., Sprehn, G., Saykin, A.J., 2002. Intact motor imagery in chronic upper limb hemiplegics: Evidence for activity-independent action representations. J. Cognit. Neurosci. 14 (6), 841–852. https://doi.org/10.1162/089892902760191072.

Konishi, Y., Kuriyama, M., Mikawa, H., Suzuki, J., 1987. Effect of body position on later postural and functional lateralities of preterm infants. Dev. Med. Child Neurol. 29 (6), 751–757.

Koo, T.K., Li, M.Y., 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J. Chiropract. Med. 15 (2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Leake, H., 2012. Investigating cortical proprioceptive maps in people with neck pain: A left/right neck rotation judgement task'. School of Physiotherapy, University of South Australia.

Lee, H., Lamb, S.E., Bagg, M.K., Toomey, E., Cashin, A.G., Moseley, G.L., 2018. Reproducible and replicable pain research: a critical reviewhttps://doi.org/10.1097/j.pain.0000000000001254. Pain, vol. Articles (in press).

Meng, S., Oi, M., Saito, G., Saito, H., 2017. The neural correlates of biomechanical constraints in hand laterality judgment task performed from other person's perspective: A near-infrared spectroscopy study. PLoS One 12 (9), e0183818. https://doi.org/10.1371/journal.pone.0183818.

Moseley, G.L., 2004. Why do people with complex regional pain syndrome take longer to recognise their affected hand? Neurology 62, 2182–2186.

Moseley, G.L., 2006. Making sense of S1 mania - are things really that simple? In: In: G L (Ed.), Topical issues in pain, vol. 5. CNS Press, Falmouth, pp. 321–340.

Moseley, G.L., Butler, D., Beams, T., Giles, T., 2012. The Graded Motor Imagery Handbook. Noigroup Publications, Adelaide.

National Center for Health Statistics, 2016. Health, United States, 2015: With Special Feature on Racial and Ethnic Health Disparities. U.S Department of Health and Human Services, Hyattsville, MD.

Neuro Orthopaedic Institute & Noigroup Publications, 2012. Recognise App. Neuro Orthopaedic Institute & Noigroup Publications, 19, North Street, Adelaide City West, South Australia 5000 AUSTRALIA.

Nico, D., Daprati, E., Rigal, F., Parsons, L., Sirigu, A., 2004. Left and right hand recognition in upper limb amputees. Brain 127 (1), 120–132. https://doi.org/10.1093/brain/awh006.

Page, S., Levine, P., Leonard, A., 2007. Mental Practice in Chronic Stroke: Results of a Randomized, Placebo-Controlled Trial. Stroke 38, 1293–1297. https://doi.org/10.1161/01.STR.0000260205.67348.2b.

Parsons, L.M., 1987. Imagined spatial transformations of one's hands and feet. Cognit. Psychol. 19 (2), 178–241. https://doi.org/10.1016/0010-0285(87)90011-9.

Portney, L., Watkins, M., 2015. Foundations of Clinical Research, 3e : Applications to

Practice. F.A. Davis, Philadelphia, PA, UNITED STATES.
Reid, E., Wallwork, S.B., Harvie, D., Chalmers, K.J., Gallace, A., Spence, C., Moseley, G.L., 2016. A new kind of spatial inattention associated with chronic limb pain? Ann. Neurol. 79 (4), 701–704. https://doi.org/10.1002/ana.24616.
Reinhart, S., Schmidt, L., Kuhn, C., Rosenthal, A., Schenk, T., Keller, I., Kerkhoff, G., 2012. Limb activation ameliorates body-related deficits in spatial neglect. Frontiers Research Foundation.
Saimpont, A., Pozzo, T., Papaxanthis, C., 2009. Aging affects the mental rotation of left and right hands. PLoS One 4 (8), e6714. https://doi.org/10.1371/journal.pone.0006714. Aug 26.
Schmid, A., Coppieters, M., 2012. Left/right judgment of body parts is selectively impaired in patients with unilateral carpal tunnel syndrome. Clin. J. Pain 28 (7), 615–622.
Sealed Envelope Ltd, 2018. *Sealed Envelope,* Randomisation and online databases for clinical trials. updated 2018, London, viewed 3 July 2018. https://www.sealedenvelope.com/.
Stanton, T.R., Lin, C.W., Smeets, R.J., Taylor, D., Law, R., Moseley, G.L., 2012. Spatially defined disruption of motor imagery performance in people with osteoarthritis. Rheumatology 51 (8), 1455–1464.
Ververs, I.A., de Vries, J.I., van Geijn, H.P., Hopkins, B., 1994. Prenatal head position from 12-38 weeks. I. Developmental aspects. Early Hum. Dev. 39 (2), 83–91 Oct 28.
von Piekartz, H., Wallwork, S.B., Mohr, G., Butler, D.S., Moseley, G.L., 2014. People with chronic facial pain perform worse than controls at a facial emotion recognition task, but it is not all about the emotion. J. Oral Rehabil. 42 (4), 243–250.
Wallwork, S.B., Butler, D.S., Fulton, I., Stewart, H., Darmawan, I., Moseley, G.L., 2013. Left/right neck rotation judgments are affected by age, gender, handedness and image rotation. Man. Ther. 18 (3), 225–230. https://doi.org/10.1016/j.math.2012.10.006.
Weir, P.J., 2005. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J. Strength Condit Res. 19 (1), 231–240.
Wong, C.K., Wong, C.K., 2017. Limb Laterality Recognition Score: A Reliable Clinical Measure Related to Phantom Limb Pain. Pain Medicine, pp. 179. https://doi.org/10.1093/pm/pnx179.
Zimney, K.J., Wassinger, C.A., Goranson, J., Kingsbury, T., Kuhn, T., Morgan, S., 2018. The reliability of card-based and tablet-based left/right judgment measurements. Musculoskeletal Sci. Pract. 33, 105–109. https://doi.org/10.1016/j.msksp.2017.09.002.
Zou, G.Y., 2012. Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. Stat. Med. 31 (29), 3972–3981. https://doi.org/10.1002/sim.5466.