Using Fitts' Law to Detect Intentional Misrepresentation

Journal of Motor Behaviour

Rebecca L Deuble, Mark J Connick, Emma M Beckman, Bruce Abernethy, and Sean M Tweedy

University of Queensland, School of Human Movement and Nutrition Sciences

Author Note

Rebecca L Deuble, University of Queensland, School of Human Movement and Nutrition Sciences; Mark J Connick, University of Queensland, School of Human Movement and Nutrition Sciences; Emma M Beckman, University of Queensland, School of Human Movement and Nutrition Sciences; Bruce Abernethy, University of Queensland, School of Human Movement and Nutrition Sciences; Sean M Tweedy, University of Queensland, School of Human Movement and Nutrition Sciences.

Correspondence concerning this article should be addressed to Rebecca Deuble, School of Human Movement and Nutrition Sciences, University of Queensland, Australia, 4072.

Tel: + 61 733654998

Fax: +61 733656877

Email: r.deuble@uq.edu.au

Abstract

In Paralympic Classification, tests of impaired coordination (e.g. reciprocal tapping tasks) are effort-dependent and therefore susceptible to Intentional Misrepresentation of Abilities (IM) - deliberate exaggeration of impairment severity. The authors investigated whether reciprocal tapping tasks performed sub-maximally could be differentiated from tapping tasks performed with maximal voluntary effort (MVE), based on conformity with Fitts' law. 10 non-disabled participants performed 14 tapping tasks with their dominant hand on three separate occasions. 7 tasks were performed with MVE and the other 7 at speeds that were at least 20% slower. Results revealed that evaluating conformity with Fitts' law is a potentially valid method for objectively detecting IM during reciprocal tapping. Evaluation of sensitivity and specificity of the method is now warranted.

Keywords: cheating, coordination, evidence-based classification, finger tapping, Paralympic Sport

Using Fitts' Law to Detect Intentional Misrepresentation

Classification in Paralympic Sport aims to minimize the impact that impairment has on competition outcomes by classifying impairments according to how much difficulty they cause in a given sport (Tweedy & Vanlandewijck, 2011). Classification systems that achieve this aim will ensure that the athletes who win are those who have best enhanced their anthropometric, physiological and psychological characteristics through legitimate means such as training and diet, rather than those who have impairments which cause less difficulty in their sport (Tweedy & Vanlandewijck, 2011).

In order to classify eligible impairments according to the extent of activity limitation they cause, requires measures of impairment that are valid, reliable, and ratio-scaled (Tweedy, Beckman, & Connick, 2014). A ratio-scaled measure of impairment is one which states unit magnitude respective to a meaningful and non-arbitrary zero value (Nunnally, 1967). In Paralympic Sport, there are eight eligible physical impairment types: hypertonia, ataxia, athetosis, limb deficiency, impaired passive range of motion, impaired muscle power, leg length difference and short stature. Of these, the first three - hypertonia, ataxia and athetosis, all adversely affect coordination.

While the validity of some measures of impairment (e.g., measures of limb deficiency or short stature) require very little participant effort/compliance, the validity of other measures of impairment (e.g. tests of strength or coordination) is highly effort-dependent. More specifically, in order to obtain a valid measure of strength or coordination, an athlete must give maximal voluntary effort (MVE) throughout. As a consequence, athletes who do not perform maximally can appear weaker or more uncoordinated than they truly are, thereby exaggerating the severity of their impairment. Athletes who engage in this behaviour have the potential to be allocated to a class for athletes who are more severely impaired, thereby increasing their chances of success.

Section 11 of the International Paralympic Committee's (IPC) Classification code defines deliberate submaximal performance as "Intentional Misrepresentation of Skills and/or Abilities" (IM) (International Paralympic Committee, 2007). There are severe punishments available to sanction athletes who engage in IM, including a lifetime ban from competition (International Paralympic Committee, 2007). However, because there are currently no objective detection methods, it is very difficult to establish that an athlete has cheated. This difficulty, together with the extensive legal and ethical consequences associated with labelling an athlete as a cheat, is likely to discourage classifiers from enforcing the rule. Methods which can objectively differentiate between MVE and IM would enable classifiers to accurately identify those who are intentionally misrepresenting their abilities and encourage them to enforce the rule.

A valid method for differentiating between MVE and IM must satisfy two main criteria: 1.There must be significant differences between the results achieved under MVE and IM conditions, and 2. Acceptable sensitivity and specificity (Portney & Watkins, 2009). Sensitivity refers to the method's ability to correctly identify participants who are intentionally misrepresenting their abilities, while specificity refers to the method's ability to correctly identify participants who are giving MVE (Portney & Watkins, 2009)

To date, research has focused on detecting IM in tests of strength within the personal injury and insurance sectors. One test that has been investigated is the Five Rung Grip test. This test required participants to exert maximal grip strength at five different handle positions, starting at the narrowest position (position 1) and ending at the widest position (position 5) (Stokes, Landrieu, Domangue, & Kunen, 1995). Stokes argued that when force of contraction was plotted against handle position, MVE performance was characterized by a skewed, bell-shaped curve, while IM performance was characterized by five uniform force

measures (Stokes, et al., 1995). However, no empirical evidence was provided to support this statement (Gutierrez & Shechtman, 2003).

Other studies found that grip, elbow flexion, and knee extension strength tests performed submaximally were characterized by a higher coefficient of variation (CV) (Robinson, Geisser, Hanson, & O'Connor, 1993; Shechtman, Anton, Kanasky, & Robinson, 2006). Unfortunately, the CV method did not have adequate sensitivity and specificity to differentiate between MVE and IM performance (Shechtman, et al., 2006). Other strength methods such as the rapid exchange grip test have also been evaluated (Shechtman & Taylor, 2000; Taylor & Shechtman, 2000; Westbrook, Tredgett, Davis, & Oni, 2002), but to date none have satisfied both essential criteria for a valid objective method (Robinson & Dannecker, 2004).

One method with potential application for detecting IM within Paralympic Sport, which has not been previously evaluated for this purpose, is Fitts' law. This law was first established in a study where participants completed a series of 15 sec reciprocal tapping tasks (Fitts, 1954). These tasks required participants to tap alternately between two target areas as quickly and as accurately as possible. The difficulty of the movement, referred to as the index of difficulty (ID) is represented by the value $\log_2 (2A/W)$, where A = the amplitude of the movement, or the distance between the two targets and W = the target width (Fitts, 1954). The ID can be increased by either decreasing W and/or by increasing A.

Fitts' study revealed that there were two established patterns of results which occurred when participants were moving as fast and as accurately as possible. The first of these was that there was a proportional increase in movement time as the ID increased (Fitts, 1954). This strong linear relationship between movement time and ID was reflected through a coefficient of determination score (R^2) that was close to one. The second characteristic described the relationship between IDs which were equivalent in difficulty, but had different

configurations of target width and amplitude. For example, a target width of 7.5 cm paired with an amplitude of 45.4 cm has the same ID as a target width of 5 cm paired with an amplitude of 30.4 cm (Fitts, 1954). According to the law, equivalent movement times should be produced by these equivalent IDs.

This preliminary study aimed to determine whether tapping tasks performed sub-maximally could be differentiated from tapping tasks performed with MVE, based on conformity with Fitts' law. In order to evaluate this, three main research questions were addressed: (1) is the R^2 score achieved significantly lower in the IM condition than in the MVE condition; (2) is the difference between the fastest and slowest times achieved for four equal but differently configured IDs significantly larger in the IM condition than in the MVE condition; and (3) is the change in the IM and the MVE results across three separate testing sessions non-significant for both (a) the R^2 score and (b) the difference between the equivalent IDs. Affirmative answers to these questions would indicate that evaluating conformity with Fitts' law is a potentially valid objective method for detecting IM on a reciprocal tapping task.

Methods

Participants

Ten non-disabled participants were recruited from the School of Human Movement Studies at the University of Queensland (five men, five women, M_{age} = 23.8 years; range 18-31 years). All participants were regularly physically active, participating in at least 3 sessions of moderate intensity physically active per week. The study was approved by the School of Human Movement Studies Ethics Committee (reference number HMS12/0914.R3). All participants gave written informed consent prior to participating. Recruitment was via promotion of the "Biggest Cheater" Competition. Posters were placed in public areas that

explained the purpose of the study and advised that there would be monetary rewards for those who were able to cheat the best ($100 = first, $50 = second, $25 = third). Details of how to win were explained within session one (see session one below).

Testing Equipment

Four pairs of custom made 17.5 cm x 12 cm fibreglass printed circuit boards (PCBs) were used. The PCBs were gold in colour with a white target area in the middle. The target width was different for each pair of PCBs: 2.5 cm, 5 cm, 7.5 cm, and 10 cm, with one pair of PCBs being used at any one time. The PCBs were designed to register contacts and comprised 60 gold plated longitudinal copper tracks. A 5 V pulse was registered as a contact when two or more consecutive tracks were bridged. Whilst in use the two PCBs were connected to a personal computer via a Musclelab unit (Ergotest, Norway; module version 8.23). This software was set to collect the data at a sampling rate of 100 Hz (Ergotest, Langesund, Norway). Movement times were continuously calculated throughout. This was defined as the time between one contact and the next.

On the underside of each PCB, four metal pegs (2.1 cm long) were attached, with one in each corner. They were designed to fit into the holes of a custom built metal frame. The PCBs were securely held within the frame, and positioned so that they were a set amplitude apart. This amplitude reflected the distance from the centre of one PCB to the centre of the other, which varied depending on the ID being performed. The IDs and their corresponding amplitudes are outlined under "Protocol".

Participant Setup

The participants sat on an adjustable stool in front of a 72.5 cm high bench with feet resting flat on top of a foot rest, 7.5 cm high. The position of the stool relative to the bench was adjusted so that once the participant was seated, the horizontal distance between the edge of

the bench and the greater trochanter of the right or left leg was 30 cm. Stool height was also

adjusted so that the height of the participant's greater trochanter was 5cm below the top

surface of the bench. The stool was then moved to the left/right to ensure that the shoulder of

the dominant limb was aligned with the midpoint between the two PCBs. The index finger of

the dominant hand was used for tapping and was dressed with a 6 cm (length) x 0.3 cm

(width) piece of copper tape with adhesive backing. This was worn to ensure contacts were

registered each time a PCB was tapped. The hand not used for tapping rested on the thigh of

the dominant leg (see Figure 1 for illustration of participant setup).

Protocol

Participants completed three one hour testing sessions, separated by one to seven days

($M \pm SD = 3.05 \pm 2.41$). This was to ensure that results were not affected by physical and/or

mental fatigue. All testing sessions took place in a private room with just one tester. Each of

the testing sessions required the participants to complete 14 reciprocal tapping tasks (7 x

MVE and 7 x IM). The seven MVE tasks and the seven IM tasks were done using the same

IDs. Three of IDs were 3 (W = 10 cm, A = 40.1 cm), 4.01 (W = 7.5 cm, A = 60.6 cm) and 4.6

(W = 2.5 cm, A = 30.4 cm). The other four IDs were all the same (ID = 3.6), but in each case

the target width and amplitude were different: 3.6a (W = 10 cm, A = 60.6 cm), 3.6b (W = 7.5

cm, A = 45.4 cm), 3.6c (W = 5.0 cm, A = 30.4 cm), 3.6d (W = 2.5 cm, A = 15.2 cm. These

will be referred to as the four equivalent IDs (See Figure 2 for the configuration of ID 3.6d).

For each ID, participants were asked to tap from one plate to the other within the

designated target areas using the index finger of their dominant hand for a period of 15 sec.

Participants were instructed to start tapping on the PCB which was on their dominant side. In

the MVE condition participants completed the seven tapping tasks "as fast and as accurately

as possible", and in the IM condition participants completed the same seven tapping tasks "at

least 20% slower than max." Two valid trials for each of the tapping tasks were required for

each condition. A trial was considered valid if 90% of the contacts were within the target area. The order the seven IDs were completed in was randomized within each condition and within each session. Random number generators were used to randomize the order of the IDs.

Session one. In the first session the 10 recruited participants were thoroughly familiarised with the concepts of evidence-based classification, IM, Fitts' law, and the requirements of the task. This included showing the participants a graph illustrating the Fitts' relationship between movement time and ID (see Appendix Figure A1). Participants were informed that they would complete the tapping task over seven randomized IDs; however they were not informed that four of these IDs were equivalent. Participants did not receive a detailed explanation of how ID was calculated, but understood that ID could be increased by decreasing the target width and/or by increasing the amplitude.

Participants were reminded that there were monetary rewards for the three persons who could cheat the best. It was explained that in order to win these rewards, all three IM attempts had to be eligible. In order for an IM attempt to be eligible, participants had to achieve IM movement times for at least two of the seven IDs which were ≥20% slower than their times achieved during the MVE condition. The speed "20% slower" was chosen based on the results from an unpublished reliability trial. This saw 20 non-disabled participants perform a reciprocal tapping task under MVE conditions with ID3.6c. Mean movement time was 0.244 sec with a standard deviation of 18% (0.044 sec), and this guided our selection of 20%. Participants who recorded three eligible IM attempts and achieved either the first, second or third highest average R^2 score for these three attempts received these monetary rewards.

In this initial testing session the participants first completed the reciprocal tapping tasks under MVE conditions followed by IM conditions. This order was chosen as it would be difficult for the participants to perform the tapping tasks under IM conditions before they

had performed them maximally. Before completing the task under IM conditions participants were shown a graph of both a successful and an unsuccessful cheating attempt (see Appendix Figures A2 and A3), and the tester demonstrated what a 20% slower speed looked like. Although it was not encouraged, in between sessions, participants were permitted to search for extra information about Fitts' law from other sources.

Before completing the trials for analysis, participants were provided with practice trials to familiarise with the reciprocal tapping task. For the MVE condition, three practice trials (slow, medium, and fast paced) were given for the hardest ID (4.6). For the IM condition, participants were also given three practice trials at the hardest ID, where they were asked to practice moving at least 20% slower than their speed in the MVE condition. Feedback on accuracy was provided during these IM practice trials, however no information on movement times was provided. In addition to these initial practice trials, the participants practiced for 5 sec each time the configuration of the ID changed.

Session two and three. At the beginning of the second and third sessions, participants were given feedback on how they performed within the IM condition of the previous session, in order to provide motivation to improve within remaining sessions. Each participant received their R^2 score and was shown a progressive competition leader board. However, participants did not receive a figure showing the linearity of their individual plot. Both sessions two and three were performed as per session one (including practice trials), except the order of the MVE and IM conditions was randomized. To ensure that each participant's MVE was a true representation of their maximal effort, the fastest mean movement time for any ID was no more than 5% slower than the fastest time achieved within the previous session (i.e., session one or two). In the case where it was more than 5% slower, participants were asked to complete another trial/s.

Post testing. After completion of the third testing session participants were sent their final IM R^2 score and their average R^2 score for the three sessions via email. This email also showed their final position on the competition leader board and the three participants with the best average R^2 score were awarded their respective prize money.

Data Sampling

The raw data files for each participant were exported into Microsoft Excel (2010), and the mean movement times for each 15 sec trial calculated. The fastest mean movement times for each of the seven IDs were used to perform a linear regression analysis for each individual. The R^2 score was then calculated for both MVE and IM conditions. This process was performed for each of the three sessions. The difference between the fastest and slowest times achieved for the four equivalent IDs (ID 3.6) was also calculated for both MVE and IM conditions. All data were exported to SPSS (version 20.0) for analysis.

Data Analysis

Two outcome variables were analysed: (a) the R^2 score, and (b) the difference between the fastest and slowest times achieved for the four equivalent IDs. In each case data were tested for normality using the Shapiro-Wilk test. A 2 x 3 (condition x session number) Repeated Measures ANOVA was performed to determine whether there were any differences between the two outcome variables under MVE and IM conditions. Paired t-tests were performed for each of these variables to determine if there were differences between IM and MVE conditions at sessions one, two and three. All P values were adjusted for multiple testing using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). An adjusted P value less than 0.05 was considered significant. A one-way Repeated Measures ANOVA was also performed to determine whether there were significant changes in the IM and MVE results across the three sessions for both outcome variables. Finally, to ensure that the instruction for the IM condition: "to move at least 20% slower than max" was understood, a 2

x 3 x7 Repeated Measures ANOVA (condition x session number x ID setup) was performed

to check that the movement times were significantly slower under IM conditions.

Results

The Shapiro-Wilk test for normality revealed that the two outcome variables were

both normally distributed. Results from the 2 x 3 (condition x session number) Repeated

Measures ANOVA revealed that the R^2 score was significantly smaller under IM conditions

compared to MVE conditions (p < 0.001); and that the difference between the equivalent IDs

was significantly greater under IM conditions compared to MVE conditions (p = 0.021).

Results from the paired t-test analyses (with Benjamini-Hochberg correction)

comparing the mean R^2 scores under MVE and IM conditions are displayed within figure 3.

The R^2 score is shown to be significantly smaller for the IM condition compared to MVE at

sessions 1 (p = 0.013), 2 (p = 0.005) and 3 (p < 0.001). Results from the paired t-test analysis,

comparing the mean differences between the four equivalent IDs under MVE and IM

conditions are displayed within figure 4. This difference is shown to be significantly greater

for the IM condition compared to MVE at sessions 1 (p = 0.038) and 3 (p = 0.006).

Results for the one-way Repeated Measures ANOVA revealed that there was no

significant change in the mean R^2 scores achieved across the three sessions under IM

conditions, F (2, 8) = 0.415, p = 0.673. There was also no significant change in the mean

differences between the four equivalent IDs under IM conditions, F (1.151, 10.356) = 1.838,

p = 0.206, with Greenhouse Geisser correction. Similarly, under MVE conditions, there was

no significant change across the three sessions in either the mean R^2 scores: F (2, 8) = 0.673,

p = 0.537, or the mean differences between the four equivalent IDs: F (2, 8) = 0.274, p =

0.767. The movement times were significantly slower for the IM condition compared to the

MVE condition at sessions 1 (p < 0.001), 2 (p < 0.001) and 3 (p < 0.001).

Discussion

The results of this investigation collectively indicate that reciprocal tapping tasks performed sub-maximally can be differentiated from tapping tasks performed with MVE, based on conformity with Fitts' law. Specifically, mean R^2 scores were significantly lower under IM conditions at all three sessions. There were also significantly greater average differences between the fastest and slowest movement times for the four equivalent IDs under IM conditions, at two of the three sessions. The third and final important observation was that greater familiarisation with the task did not significantly improve MVE or IM scores for either outcome variable. The results therefore indicate that evaluating conformity with Fitts' law is a potentially valid method for objectively detecting IM on a reciprocal tapping task.

Research efforts to develop an objective method are not only important because they will permit detection of IM, but because according to the Valency Instrumentality and Expectancy (VIE) theory, objective methods will also act to reduce an athlete's motivation to cheat (Ellingson & McFarland, 2011). The VIE theory indicates that an athlete's motivation to intentionally misrepresent their abilities during classification will be determined by three factors: (a) valence - an athlete believes that achieving competitive success by IM will bring personal satisfaction. Valence is increased by increasing kudos and monetary reward for Paralympic success; (b) instrumentality - an athlete believes that IM is critical for achieving competitive success. Instrumentality is increased when methods for detecting IM are subjective and decreased when methods are objective; and (c) expectancy - when an athlete feels confident in their ability to cheat successfully and achieve favourable classification. Expectancy will be reduced if athletes are warned that classification included objective methods for detecting IM (Ellingson & McFarland, 2011). According to VIE theory, development of objective methods for detecting IM will reduce athlete motivation to intentionally misrepresent by decreasing both instrumentality and expectancy.

Findings from this study are consistent with those of two previous studies which concluded that Fitts law is violated under submaximal conditions (Maruff & Velakoulis, 2000; Young, Pratt, & Chau, 2009). However, the present study is a critical step forward because the experimental design was much more rigorous than these previous studies. Specifically: participants were familiarised with the purpose of the task or Fitts' law; participants were given multiple attempts under both IM and MVE conditions; and participants were motivated to try as hard as possible because there was a monetary reward for those who were best able to conform to Fitts' law. One minor limitation of the design is that the number of days between the sessions was variable ($M \pm SD = 3.05 \pm 2.41$), providing some participants with more opportunity to develop potential cheating strategies than others. Future studies should therefore consider standardising the number of days between sessions. Overall however; the rigor of the design used in this study indicates that Fitts' fundamental law of movement speed and accuracy is more robust than has previously been demonstrated (Maruff & Velakoulis, 2000; Young, et al., 2009).

One other important difference in the current study design is the method of calculation used for the R^2 score. In the current study the mean R^2 scores reported are lower than those reported in Fitts' related studies (Fitts, 1954; Fitts & Peterson, 1964). This is because in Fitts original studies, the main outcome measure was the mean movement times achieved by the group for each ID. Linear regression was then used to report an R^2 score for the group (Fitts & Peterson, 1964). In contrast, in the current study, R^2 scores were calculated for each individual participant, and used as an indicator of whether an individual was giving MVE or intentionally misrepresenting their abilities. The mean R^2 score reported in the current study therefore reflected the average of the individual R^2 scores achieved by the ten participants. When data from the first testing session were re-analysed using a regression

model of the group's mean movement times, the R^2 score was 0.95. This is comparable to the 0.99 R^2 score reported by Fitts (Fitts & Peterson, 1964).

Although the results from the group level analysis indicate that the method is potentially valid, the large standard deviations under IM conditions suggest that some participants would avoid detection (see Figures 3 and 4). This is emphasized by the large range of scores achieved for both outcome variables under IM conditions. For example, in session 1 the range of R^2 scores was 0.009 to 0.913, while the range of differences between the equivalent IDs was 0.029-0.338. However, what is promising is that most participants did not achieve their best scores under IM conditions during session three when most familiarised with the task. For example, the R^2 score achieved by 70% of participants in session three, was lower than the scores achieved in either session one or two. Similarly, the difference between the equivalent IDs achieved by 80% of participants in session three was greater than the difference achieved in either session one or two. These results provide further evidence to suggest that there was no significant improvement in the IM results with greater task familiarisation.

The results from this study are positive; however in order to satisfy the second criterion for a valid IM detection method, more specific individual level analyses are required. Future research should determine whether these two outcome variables have sufficient sensitivity and specificity to be validly applied in isolation. If sensitivity and specificity values are insufficient, then there is also the potential to use these two features in combination, or with other data features. For example, spatial features such as the variability of the horizontal movement amplitude over a given trial, could be combined with Fitts' law to improve sensitivity and specificity outcomes. Combining spatial features with Fitts' law may potentially improve the method's accuracy and validity for detecting IM.

In order to evaluate the sensitivity and specificity of these methods, the first step is to determine an appropriate sample size for subsequent studies, which is informed by the results obtained within the current study. These future studies should include more sophisticated analyses which will determine sensitivity and specificity firstly amongst non-disabled participants. For example, a Reciever Operating Characteristic (ROC) Curve analysis can be used to find an optimal cut-off value which yields the best combination of sensitivity and specificity (MacNicol, 2005). High levels of sensitivity will reduce the likelihood of a false negative result (i.e., an athlete engaging in IM who is not identified), while high levels of specificity will reduce the likelihood of a false positive result (i.e., an athlete being wrongly identified for IM) (Ghori & Chung, 2007).

In the context of classification, greater priority should be placed on specificity; as if an athlete is wrongly accused this would be extremely detrimental to their career, prize money and self-esteem. It would also potentially expose the Paralympic Movement to legal action, as an athlete found to not be cheating would have a strong case. However, sensitivity is still an important secondary consideration because the method needs to be able to effectively identify the cheats.

If levels of sensitivity and specificity are deemed to be sufficient amongst non-disabled participants, further studies should then establish whether the same results can be applied to a sample of athletes with coordination impairments (i.e. those with hypertonia, ataxia and/or athetosis). Fitts' law has been shown to hold amongst individuals with impaired coordination (Bertucco & Sanger, 2014; Smits-Engelsman, Rameckers, & Duysens, 2007). However, in order to be usefully applied to detect IM within Paralympic Sport, the findings of the current study need to be replicated in athletes with coordination impairments. These follow up studies will be a crucial step in assessing whether the method has adequate sensitivity and specificity to detect athletes who intentionally misrepresent their abilities

within Paralympic Sport. These results may be applied to detect IM in the assessment of compensable injury, specifically those with Traumatic Brain Injury. There is also the potential to explore other IM methods, such as strength tests, which can be applied to detect those who intentionally appear weaker than they truly are.

In summary, the results of this pilot study collectively suggest that evaluating conformity with Fitts' law is a potentially valid method for detecting IM in a reciprocal tapping task. Although the results are promising, future studies are now warranted to determine whether the R^2 score and/or the difference between the equivalent IDs have sufficient sensitivity and specificity to detect IM. There is also the potential to combine Fitts' law with spatial features associated with the task in order to improve the method's accuracy. To establish whether these features can be combined to detect IM in a reciprocal tapping task, sufficiently powered ROC curve analyses will be required to estimate sensitivity and specificity. Studies must also assess whether the results from this study can be replicated in athletes with coordination impairments. Objective tests with acceptable sensitivity and specificity are important within Paralympic Sport, as they will not only help to detect IM but also reduce instrumentality and expectancy.

Reference List

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological, 57*(1), 289-300.

Bertucco, M., & Sanger, T. D. (2014). Speed-accuracy testing on the apple iPad(R) provides a quantitative test of upper extremity motor performance in children with dystonia. *Journal of Child Neurology, 29*(11), 1460-1466. doi: 10.1177/0883073813494265

Ellingson, J. E., & McFarland, L. A. (2011). Understanding faking behavior through the lens of motivation: an application of VIE theory. *Human Performance, 24*(4), 322-337. doi: 10.1080/08959285.2011.597477

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381-391.

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67*, 103-112.

Ghori, A. K., & Chung, K. C. (2007). A decision-analysis model to diagnose feigned hand weakness. *Journal of Hand Surgery-American Volume, 32A*(10), 1638-1643. doi: 10.1016/j.jhsa.2007.09.010

Gutierrez, Z., & Shechtman, O. (2003). Effectiveness of the five-handle position grip strength test in detecting sincerity of effort in men and women. *American Journal of Physical Medicine & Rehabilitation, 82*(11), 847-855. doi: 10.1097/01.phm.0000083667.25092.4e

International Paralympic Committee. (2007). International Paralympic Committee (IPC) classification code and international standards. Retrieved 08/09/2014 http://www.paralympic.org/sites/default/files/document/120201084329386_2008_2_classification_code6.pdf

MacNicol, D. A. (2005). *Primer of signal detection theory*. London, UK: Lawrence Erlbaum

Associates, Publishers.

Maruff, P., & Velakoulis, D. (2000). The voluntary control of motor imagery. Imagined

movements in individuals with feigned motor impairment and conversion disorder.

Neuropsychologia, 38(9), 1251-1260. doi: 10.1016/S0028-3932(00)00031-2

Nunnally, J. C. (1967). *Psychometric theory*. New York: McGraw-Hill.

Portney, L. G., & Watkins, M. P. (2009). *Foundations of clinical research, applications to*

practice. Norwalk, CT: Appleton and Lange

Robinson, M. E., & Dannecker, E. A. (2004). Critical issues in the use of muscle testing for

the determination of sincerity of effort. *Clinical Journal of Pain, 20*(6), 392-398. doi:

10.1097/00002508-200411000-00003

Robinson, M. E., Geisser, M. E., Hanson, C. S., & O'Connor, P. D. (1993). Detecting

submaximal efforts in grip strength testing with the coefficient of variation. *Journal of*

Occupational Rehabilitation, 3(1), 45-50.

Shechtman, O., Anton, S. D., Kanasky, W. F., Jr., & Robinson, M. E. (2006). The use of the

coefficient of variation in detecting sincerity of effort: a meta-analysis. *Work, 26*(4),

335-341.

Shechtman, O., & Taylor, C. (2000). The use of the rapid exchange grip test in detecting

sincerity of effort, part II: validity of the test. *Journal of Hand Therapy, 13*(3), 203-

210.

Smits-Engelsman, B. C. M., Rameckers, E. A. A., & Duysens, J. (2007). Children with

congenital spastic hemiplegia obey Fitts' law in a visually guided tapping task.

Experimental Brain Research, 177(4), 431-439. doi: 10.1007/s00221-006-0698-x

Stokes, H. M., Landrieu, K. W., Domangue, B., & Kunen, S. (1995). Identification of low-effort patients through dynamometry. *Journal of Hand Surgery-American Volume, 20A*(6), 1047-1056.

Taylor, C., & Shechtman, O. (2000). The use of the rapid exchange grip test in detecting sincerity of effort, part I: administration of the test. *Journal of Hand Therapy, 13*(3), 195-202.

Tweedy, S. M., Beckman, E. M., & Connick, M. J. (2014). Paralympic classification: conceptual basis, current methods, and research update. *Physical Medicine and Rehabilitiation, 6*(8 Suppl), S11-17. doi: 10.1016/j.pmrj.2014.04.013

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*(4), 259-269. doi: 10.1136/bjsm.2009.065060

Westbrook, A. P., Tredgett, M. W., Davis, T. R. C., & Oni, J. A. (2002). The rapid exchange grip strength test and the detection of submaximal grip effort. *Journal of Hand Surgery-American 27A*(2), 329-333. doi: 10.1053/jhsu.2002.30910

Young, S. J., Pratt, J., & Chau, T. (2009). Target-directed movements at a comfortable pace: movement duration and Fitts's law. *Journal of Motor Behavior, 41*(4), 339-346. doi: 10.3200/JMBR.41.4.339-346

Figure 1. Participant positioning for reciprocal tapping task (RTT). In the top right corner, a close up of the tapping pad worn on the index finger is shown.
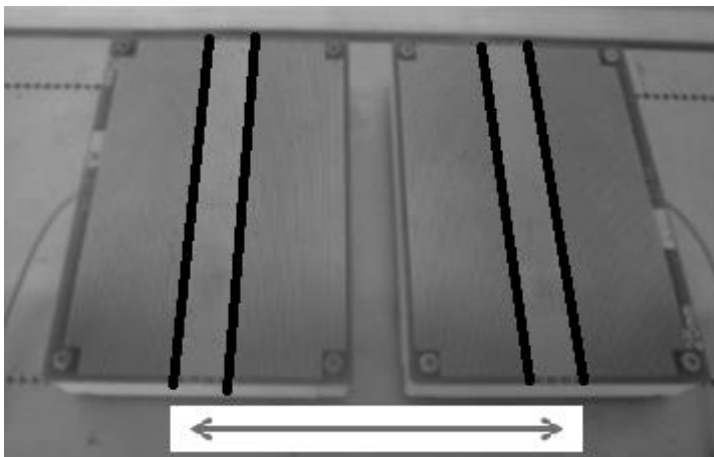

Figure 2. One of the seven combinations of target width and amplitude used. Index of difficulty (ID) = 3.6d (amplitude = 15.2cm, width = 2.5cm). The amplitude reflects the distance from the middle of one target area to the middle of the other target area (as indicated by the grey arrow). The target area/width is indicated by the two black lines in the middle of the two plates.
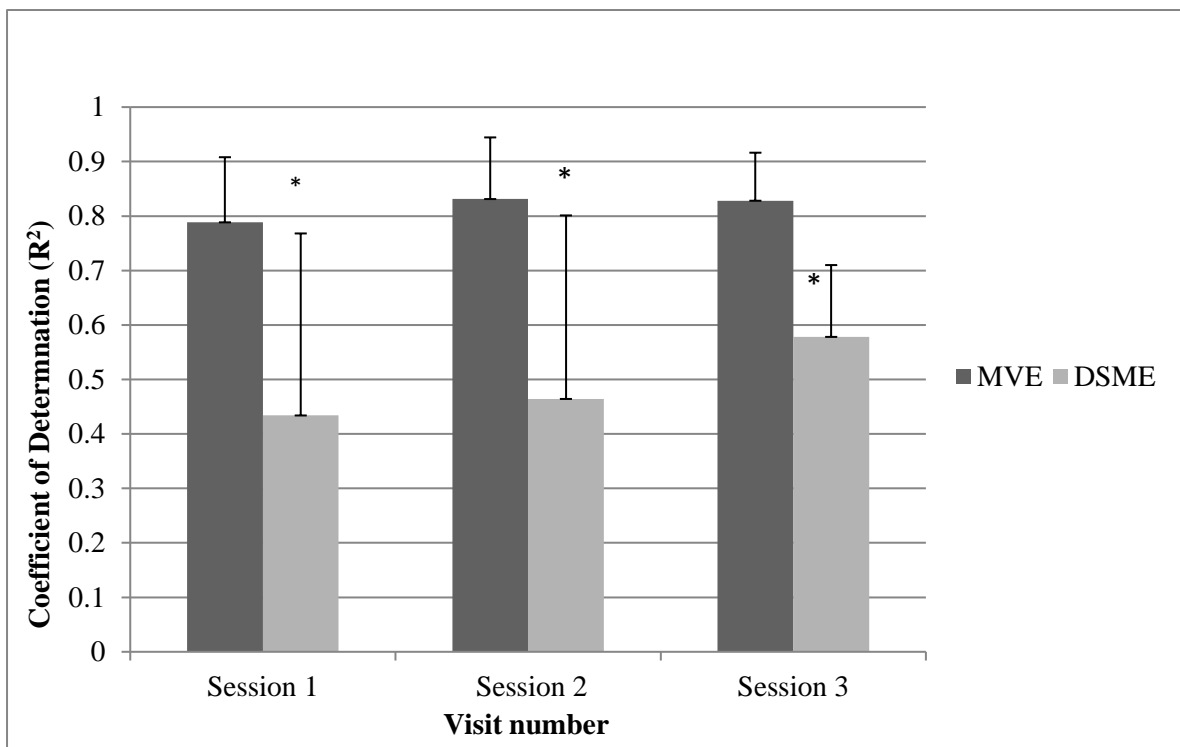
Figure 3. Mean R^2 values (+SD) achieved for both Maximum Voluntary Effort (MVE) (dark bars) and Deceptive Submaximal Effort (DSME) conditions (light bars) for sessions 1, 2 and 3. (* $p < 0.05$)
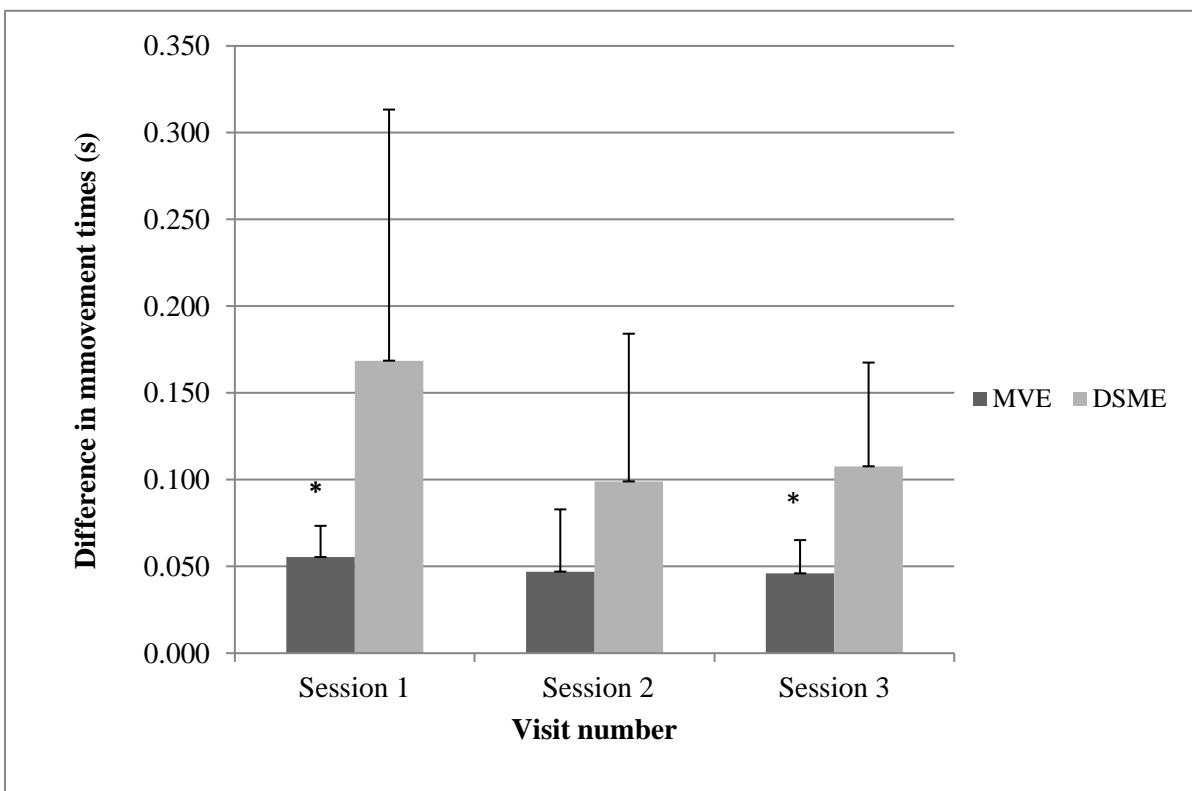

Figure 4. Mean differences (+SD) between the shortest and longest movement times for four differently configured Index of Difficulties (IDs): 3.6 (a, b, c, d), for Maximum Voluntary Effort (MVE) (dark bars) and Deceptive Submaximal Effort (DSME) conditions (light bars) at sessions 1, 2 and 3. (* $p < 0.05$)
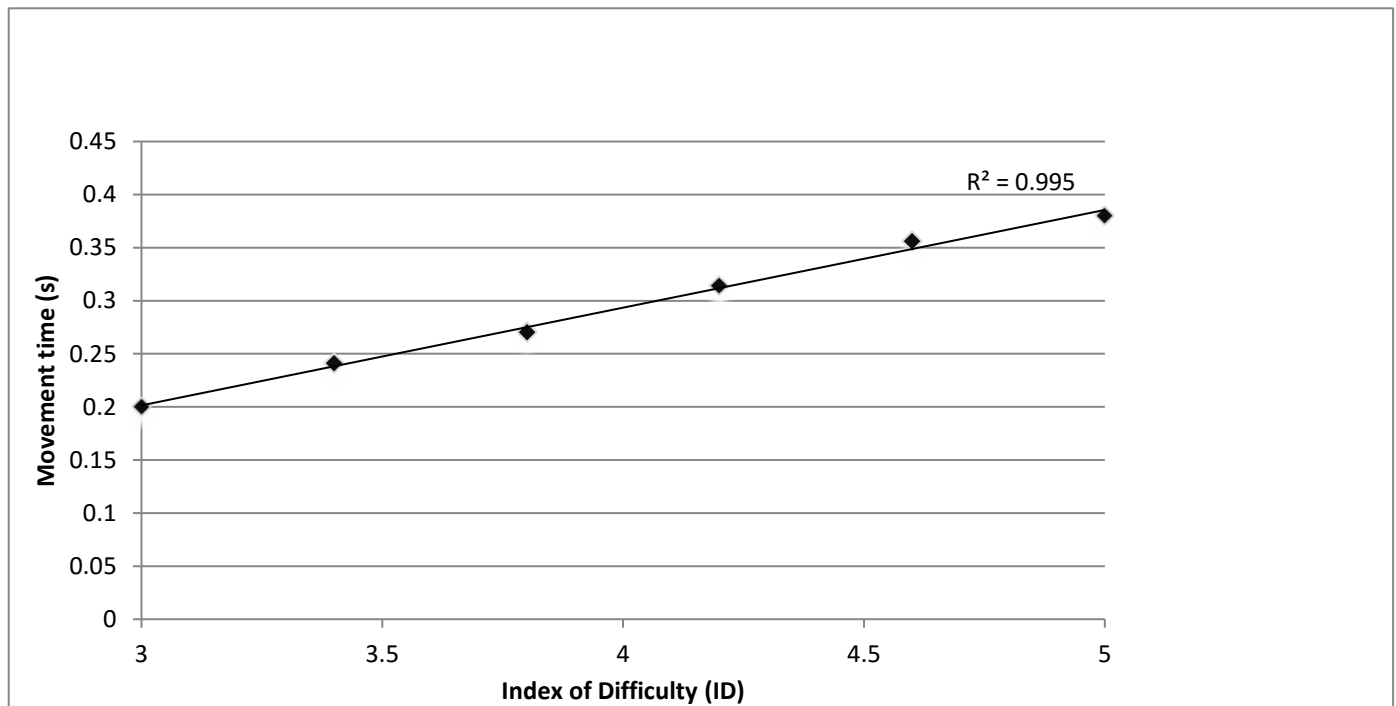
Appendix: Fitts' law visuals

Figure A1. Graph showing the pattern of results (straight line) that are expected, provided the participant is moving as fast and as accurately for each trial.
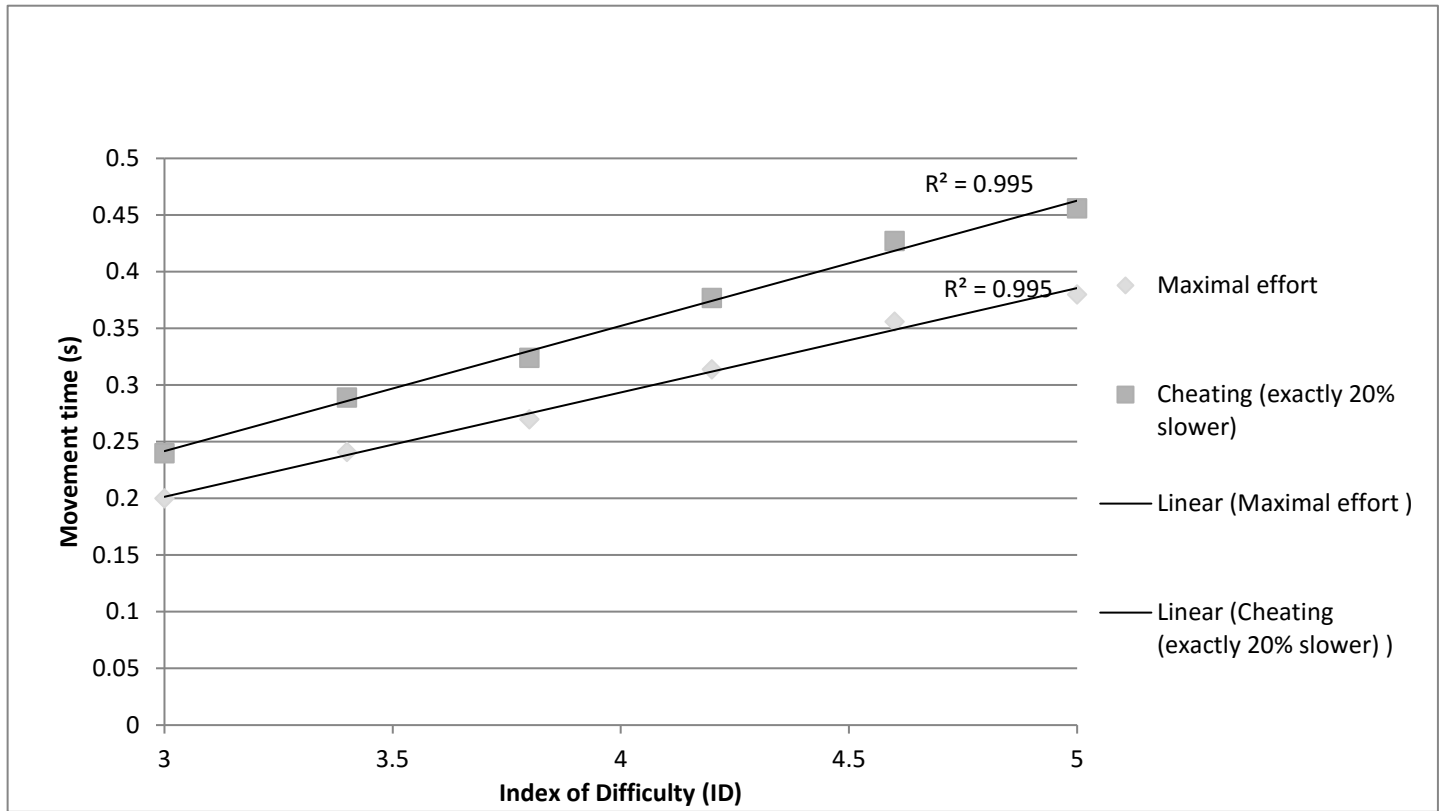
Figure A2. Graph showing a successful cheating attempt. The line with the boxes shows a person moving as fast and as accurately as possible, while the line with the diamonds shows the same person moving exactly 20% slower for each level of difficulty. These points are all in a straight line and the R^2 values are identical, therefore indicating that they were successful in their cheating attempt.
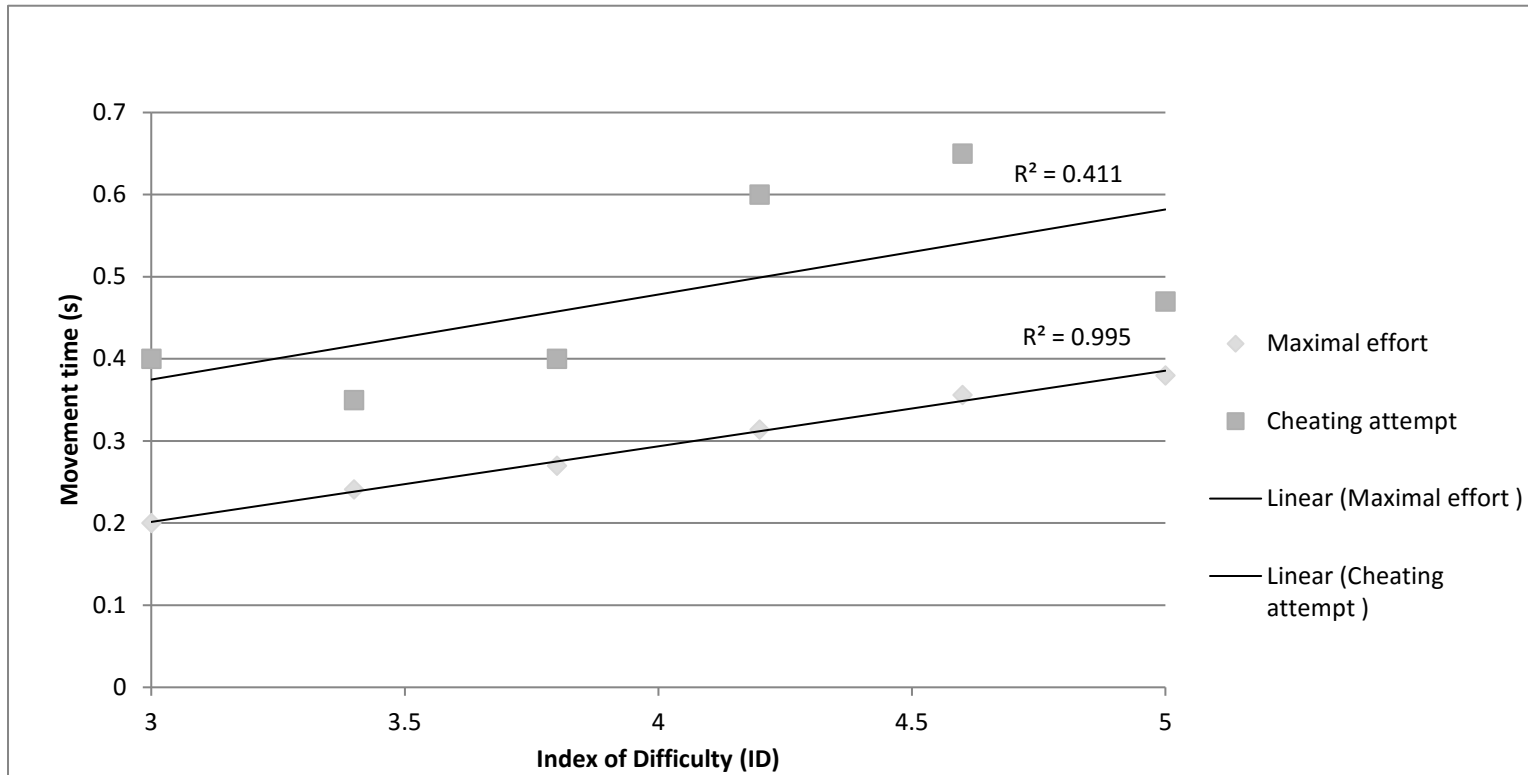
Figure A3: Graph showing an unsuccessful cheating attempt. The line with the boxes shows
a person moving as fast and as accurately as possible. The line with the diamonds shows the
same person moving at speeds that are slower. However these points are not on a straight
line, and the R^2 value is a significantly lower than that achieved for the line with boxes. This
therefore indicates that they were unsuccessful in their cheating attempt.