Full title: Peak oxygen uptake measured during a perceptually-regulated exercise test is reliable in community-based manual wheelchair users.

Running title: Perceptually-regulated exercise testing in manual wheelchair users.

Authors: Michael J. Hutchinson (m.j.hutchinson@lboro.ac.uk)[a], Maureen J. MacDonald (macdonmj@mcmaster.ca)[a,b], Roger Eston (roger.eston@unisa.edu.au)[c], Victoria L. Goosey-Tolfrey (v.l.tolfrey@lboro.ac.uk)[a].

Author affiliations: [a] The Peter Harrison Centre for Disability Sport, School for Sport, Exercise and Health Sciences, Loughborough University, LE11 3TU, United Kingdom; +44 (0)1509 226387.

[b] Exercise Metabolism Research Group, Department of Kinesiology, McMaster University, Hamilton, ON L8S 4L8, Canada; +1 905 525 9140 (ext. 22616).

[c] Alliance for Research in Exercise, Nutrition and Activity, Sansom Institute for Health Research, School of Health Sciences, University of South Australia, Adelaide SA 5000, Australia; +61 8 8302 2245.

Corresponding author: Professor Victoria L. Goosey-Tolfrey[a].

Abstract

This study aimed to compare test-retest reliability and peak exercise responses from ramp-incremented (RAMP) and maximal perceptually-regulated ($PRET_{max}$) exercise tests during arm crank exercise in individuals reliant on manual wheelchair propulsion (MWP). Ten untrained participants (9 male) completed four trials over a 2-week period, performing two RAMP (0-40 W + 5-10 $W \cdot min^{-1}$) trials one week followed by two $PRET_{max}$ trials the next, or vice versa. $PRET_{max}$ consisted of five, 2-min stages performed at Ratings of Perceived Exertion (RPE) 11, 13, 15, 17 and 20. Participants freely changed the power output to match the required RPE. Gas exchange variables, heart rate, power output, RPE and affect were determined throughout trials. The $\dot{V}O_{2peak}$ from RAMP (14.8 ± 5.5 $ml \cdot kg^{-1} \cdot min^{-1}$) and $PRET_{max}$ (13.9 ± 5.2 $ml \cdot kg^{-1} \cdot min^{-1}$) trials were not different ($P = 0.08$). Measurement error was 1.7 and 2.2 $ml \cdot kg^{-1} \cdot min^{-1}$ and coefficient of variation 5.9% and 8.1% for measuring $\dot{V}O_{2peak}$ from RAMP and $PRET_{max}$, respectively. Affect was more positive at RPE 13 ($P = 0.02$), 15 ($P = 0.01$) and 17 ($P = 0.01$) during $PRET_{max}$. This study shows the $PRET_{max}$ can be used to measure $\dot{V}O_{2peak}$ in participants reliant on MWP and leads to a more positive affective response compared to RAMP.

Key words: oxygen consumption; RPE; disability; exercise testing; test-retest reliability

Introduction

Important determinants of physical capacity in individuals dependant on manual wheelchair propulsion (MWP) include unmodifiable factors, such as age, gender and type of disability. Yet sports participation (Janssen, Dallmeijer, Veeger, & van der Woude, 2002) and exercise training (Hicks et al., 2011; Valent, Dallmeijer, Houdijk, Talsma, & van der Woude, 2007) can positively affect physical capacity in persons reliant on MWP. Furthermore, increased physical capacity, as measured using peak oxygen uptake ($\dot{V}O_{2peak}$) is linked with improved physical functional status (Dallmeijer & van der Woude, 2001), life satisfaction (van Koppenhagen et al., 2014) and self-independency during activities of daily living (Hjeltnes & Jansen, 1990) in this population. Considering these known benefits of increased $\dot{V}O_{2peak}$, as well as increasing life expectancy for people reliant on MWP (Middleton et al., 2012; Savic et al., 2017), there is a need for appropriate protocols with which to measure $\dot{V}O_{2peak}$.

Traditionally, ramp-incremented (RAMP) tests which feature fixed increases in power output (PO) that continue until volitional exhaustion (Whipp, Davis, Torres, & Wasserman, 1981) have been adopted for both able-bodied and disability groups. However, despite a RAMP protocol being the most common method for the direct measurement of $\dot{V}O_{2peak}$, this form of exercise testing has, in recent years, been subject to criticism in that it is ‘open loop’ in nature, i.e., it has no predetermined or known end-point, and therefore does not allow for pacing to occur (Noakes, 2008). An alternative to RAMP testing which is gaining in popularity in the scientific literature, and recently described as a paradigm shift in exercise testing methodology (Beltz et al., 2016) is to progress the intensity based on incremental clamping of the Ratings of Perceived Exertion (RPE), as opposed to PO.

Research has validated the use of a maximal perceptually-regulated exercise test ($PRET_{max}$) to measure $\dot{V}O_{2peak}$ during cycle (Straub, Midgley, Zavorsky, & Hillman, 2014) and handcycle exercise (Hutchinson, Paulson, Eston, & Goosey-Tolfrey, 2017) against RAMP protocols. However, the PRETmax method has yet to be applied to participants reliant on MWP for daily activity. The $PRET_{max}$ consists of five 2-min stages clamped at RPE 11, 13, 15, 17 and 20 on Borg's 6-20 RPE scale (Borg, 1998). Importantly, the $PRET_{max}$ is of fixed duration and allows the participant to control the workload and pacing strategy, satisfying the major criticisms of the RAMP protocol.

The use of $PRET_{max}$ in exercise testing of participants reliant on MWP may be justified when considering the affective response to exercise. Previous research in able-bodied participants has shown that exercise at a self-selected intensity, as in the $PRET_{max,}$ leads to a more positive affective response compared to imposed exercise of the same intensity, as in the RAMP (Evans, Parfitt, & Eston, 2014; Hamlyn-Williams, Freeman, & Parfitt, 2014; Rose & Parfitt, 2007) Hence, the $PRET_{max}$ may be a preferred option to use instead of RAMP, particularly for older participants, or those who are beginning to become more physically active.

This is the first study to investigate the use of a $PRET_{max}$ in a population with a disability. The aim of this study was to assess the reliability of the $PRET_{max}$ to measure peak exercise responses in participants reliant on MWP and to compare the responses between $PRET_{max}$ and RAMP. A further aim was to investigate the affective response to $PRET_{max}$ and RAMP protocols. It was hypothesised that the $PRET_{max}$ and RAMP would produce similar maximal exercise responses, and that affect would be more positive during $PRET_{max}$ than RAMP.

Methods

Participants

Ten (9 male, 1 female), sedentary or recreationally active MWP participants gave written informed consent to participate in this study, which was approved by the Hamilton Health Sciences Integrated Research Ethics Board (Ref. #1615). Descriptive characteristics are presented in Table 1, which represent a group typical of that commencing an exercise program as part of the Physical Activity Centre of Excellence at McMaster university. Participants were deemed safe and appropriate to take part in this study as a result of being cleared by a physician prior to joining the exercise program. This included having completed a maximal exercise test.

****Table 1 near here****

Experimental design

Following a randomised, crossover design, participants completed four maximal exercise tests over a two-week period (Figure 1) while seated in their everyday wheelchair. Trials were separated by 48 to 96 hours. All testing was conducted using the same wall-mounted electrically braked arm crank ergometer (Lode Angio, Lode B. V., Groningen, Netherlands) operating asynchronously. The ergometer was adjusted so that the centre of the crank axis was level with the shoulder and so there was slight elbow flexion at the furthest point of the crank cycle.

****Figure 1 near here****

All trials were performed at the same time of day within each participant to minimise diurnal variations (Hill, Cureton, & Collins, 1989) and dietary intake was replicated in the 24 hours before all trials. Participants refrained from alcohol consumption and vigorous exercise for 24 hours, and caffeine for 6 hours preceding each trial. Participants preferred cadence was established in the warm-up to their first trial, when they were invited to experiment with various cadences and choose what they preferred. This cadence was then recorded and subsequently participants were asked to maintain it at that level for the maximal trials.

Ramp-incremented $\dot{V}O_{2peak}$ test (RAMP) and verification stage (VER)

The RAMP started at 0-40 W and was increased by 5-10 W·min^{-1} until volitional exhaustion or preferred cadence could not be maintained. Starting PO and the PO increment were individualised for participants to match the RAMP test duration to that of the PRET$_{max}$ (10 min). Gas exchange variables were collected throughout using a facemask (7450 Series V2, Hans Rudolph Inc., Shawnee, USA) and online gas analysis system (Moxus Metabolic System, AEI Technologies Inc., Pittsburgh, USA). Heart rate (HR) was assessed throughout (RS400, Polar, Kempele, Finland) and differentiated measures of peripheral (RPE$_P$), central (RPE$_C$) and overall (RPE$_O$) RPE (Borg, 1998) as well as Feeling Scale (FS) rating (Hardy & Rejeski, 1989) were verbally recalled in the final 15 s of each stage. The FS ranges from +5 (very good) to -5 (very bad) with anchors at +3 (good), +1 (fairly good), 0 (neutral), -1 (fairly bad) and -3 (bad). Prior to all trials participants were read standardised instructions on the use of Borg’s 6-20 RPE scale (Borg, 1998).

Following termination of the RAMP participants completed 10 min of recovery (unloaded arm cranking and/or seated rest) before performing the verification phase (VER). PO was increased by 5 W from the end of the RAMP and participants cranked again until volitional exhaustion or cadence could not be maintained. Gas exchange variables, HR and subjective measures were recorded as during RAMP. Throughout RAMP and VER participants maintained their preferred cadence, which along with the subjective scales, was the only information visible to participants.

Perceptually regulated $\dot{V}O_{2peak}$ test ($PRET_{max}$)

During $PRET_{max}$ participants completed five, two-minute stages where RPE_O was clamped and progressively increased with each stage. Stages corresponded to RPE 11, 13, 15, 17 and 20 on Borg's 6-20 RPE scale (Borg, 1998). Participants self-regulated the PO by saying "up" or "down", where the investigator would adjust the PO by 3 W accordingly. Participants were not aware of the magnitude of the change but were instructed to change PO as often as required to maintain the desired RPE and to reach maximal exertion at the end of the final stage. As with during RAMP, participants maintained their preferred cadence throughout and had cadence along with subjective scales in their line of sight. Elapsed time was also visible during $PRET_{max}$ to allow pacing in relation to the end point of the exercise bout. $\dot{V}O_2$, HR and subjective measures were recorded as they were during RAMP.

Data Processing and Statistical Analysis

Analysis was performed using IBM SPSS Statistics 22 (SPSS Inc., Chicago, IL.). Physiological data are presented as mean ± SD, whilst subjective data are presented as median (interquartile range). Statistical significance was accepted at $P < 0.05$. For all tests HR and gas exchange variables were subjected to a 30 s rolling average with the highest value taken as the peak response. During $PRET_{max}$ PO was also subjected to a 30 s rolling average with the highest value taken as the peak PO (PO_{peak}). For RAMP trials PO_{peak} was calculated based on the final completed stage and proportion of the next stage completed using the formula:

$$\mathrm{POpeak} = \mathrm{F} + \left(\left(\frac{\mathrm{t}}{60\mathrm{s}}\right) \times \mathrm{I}\right).$$

Where F = PO of the final completed stage, t = time spent in the final, uncompleted stage in seconds, 60 s = stage duration and I = the PO increment. In keeping with the assessment of maximal exercise responses, the greater responses for RAMP and $PRET_{max}$ from repeat tests were used in subsequent analysis.

Reliability of peak physiological variables was assessed by calculating the coefficient of variation , and the intraclass correlation coefficient ($ICC_{3,1}$) using an openly available spreadsheet (Hopkins, 2015). The $ICC_{3,1}$ were interpreted for their magnitude in accordance with Munro's criteria where 0-0.25 is "little to no" correlation, 0.26-0.49 "low" correlation, 0.50-0.69 "moderate" correlation, 0.70-0.89 "high correlation" and 0.90-1.00 "very high" correlation (Plichta, Kelvin, & Munro, 2013). Furthermore, measurement error (ME) was calculated as the within-subject standard deviation and the smallest detectable difference (SDD) as 2.77 multiplied by ME (Bland & Altman, 1996).

Data were checked for normal distribution using the Shapiro-Wilk test statistic. Familiarisation with peak exercise testing across trial 1 to 4 was investigated using one-way repeated measures Analysis of Variance (ANOVA). Differences in test duration and peak physiological responses between protocols were assessed via paired samples t-test and for maximal subjective responses using Wilcoxon Signed Rank test. Bland-Altman plots with 95% limits of agreement (LoA) were produced to assess the agreement for $\dot{V}O_{2peak}$, HR_{peak} and peak respiratory exchange ratio (RER_{peak}) between the two protocols (Bland & Altman, 1999).

Individual RPE:$\dot{V}O_2$, RPE:HR and RPE:PO linear relationships were determined for RPE_P, RPE_C and RPE_O during RAMP and $PRET_{max}$. These relationships underwent a Fisher transformation to allow the calculation of group averages. Differences in group correlations were assessed by two-way Analysis of Variance (ANOVA), with repeated measures on protocol (RAMP x $PRET_{max}$) and mode of RPE (RPE_P x RPE_C x RPE_O). FS ratings were extracted from the RAMP and $PRET_{max}$ corresponding to RPE 11, 13, 15 and 17, or by interpolation if the specific RPE was not reported. Wilcoxon Signed Rank test was then used to assess difference in FS rating between protocols at each RPE value.

A priori power analysis was conducted in G*Power 3.1 using the test-retest reliability statistics for absolute $\dot{V}O_{2peak}$ from a previous study involving individuals with a spinal cord injury performing wheelchair ergometry (Leicht, Tolfrey, Lenton, Bishop, & Goosey-Tolfrey, 2013). For statistical power of 0.80 and α equal to 5%, it was deemed that 10 participants would be required to find a significant difference between RAMP and $PRET_{max}$.

Results

Each participant completed all four trials and there were no missing data points. ANOVA revealed no learning effect as no significant differences were found across trial one to trial four for absolute $\dot{V}O_{2peak}$ ($F_{2.1} = 0.343$, $P = 0.73$), relative $\dot{V}O_{2peak}$ ($F_{2.1} = 0.402$, $P = 0.65$), HR_{peak} ($F_{1.5} = 2.314$, $P = 0.14$) or PO_{peak} ($F_{1.7} = 0.328$, $P = 0.69$). There was no significant difference between RAMP and VER for absolute $\dot{V}O_{2peak}$ (1.3 ± 0.3 versus 1.3 ± 0.3 $L \cdot min^{-1}$, $t_{17} = -0.441$, $P = 0.67$), relative $\dot{V}O_{2peak}$ (14.1 ± 4.3 versus 14.7 ± 4.7 $ml \cdot kg^{-1} \cdot min^{-1}$, $t_{17} = -0.747$, $P = 0.47$) or HR_{peak} (139 ± 27 versus 135 ± 27 $beats \cdot min^{-1}$, $t_{18} = 1.108$, $P = 0.28$).

Reliability

Test-retest reliability for peak physiological variables obtained from RAMP and $PRET_{max}$ are shown in Table 2. The $ICC_{3,1}$ was classified as “very high” for absolute and relative $\dot{V}O_{2peak}$, HR_{peak}, PO_{peak} and RER_{peak} for both RAMP and $PRET_{max}$.

****Table 2 near here****

Agreement between protocols

Peak responses for RAMP and $PRET_{max}$ are presented in Table 3. The Pearson correlation between responses from RAMP and $PRET_{max}$ was $r = 0.922$ ($P < 0.05$) and $r = 0.969$ ($P < 0.05$) for absolute and relative $\dot{V}O_{2peak}$, respectively. Bland-Altman plots with 95% LoA for absolute and relative $\dot{V}O_{2peak}$, HR_{peak} and RER_{peak} are displayed in Fig. 2.

****Table 3 near here****

****Figure 2 near here****

Group averaged correlations are shown in Table 4. For the RPE:$\dot{V}O_2$ relationship there was no effect of protocol ($F_{(1.0)} = 0.002$, $P = 0.96$) or mode of RPE ($F_{(1.1)} = 0.127$, P 0.75). Similarly for RPE:HR there was no effect of protocol ($F_{(1.0)} = 0.150$, $P = 0.71$) or mode of RPE ($F_{(1.4)} = 1.362$, $P = 0.28$). For the RPE:PO relationship there was a significant effect of protocol ($F_{(1.0)} = 8.025$, $P = 0.02$), with Bonferroni post-hoc comparison showing that the relationship was stronger in RAMP compared to $PRET_{max}$. There was no effect of mode of RPE on the RPE:PO relationship ($F_{(1.2)} = 0.968$, $P = 0.36$).

****Table 4 near here****

Affective response

The peak FS rating was significantly smaller during RAMP (Z = -2.368, P = 0.02) compared to $PRET_{max}$ (Table 2). There was no significant difference in average FS rating between protocols (Z = -1.265, P = 0.21). At submaximal RPE values, affect was significantly more positive during $PRET_{max}$ compared to RAMP at RPE 13 (Z = -2.403, P = 0.02), 15 (Z = -2.539. P = 0.01) and 17 (Z = -2.527, P = 0.01), see Table 5.

****Table 5 near here****

Discussion

The main finding of this study was that there was no significant difference in $\dot{V}O_{2peak}$ between RAMP and $PRET_{max}$. Furthermore, the measurement error (ME) for measuring $\dot{V}O_{2peak}$ using $PRET_{max}$ or RAMP was greater than the mean difference in $\dot{V}O_{2peak}$ between protocols. Therefore, these findings support the use of $PRET_{max}$ for measuring $\dot{V}O_{2peak}$ in participants reliant on MWP. The finding of similar $\dot{V}O_{2peak}$ values between $PRET_{max}$ and RAMP corroborates findings from research involving able-bodied participants performing lower (Chidnok et al., 2013; Evans et al., 2014; Hanson et al., 2016; Hanson, Reid, Cornwell, Lee, & Scheadler, 2017; Lim, Lambrick, Mauger, Woolley, & Faulkner, 2016; Straub et al., 2014) and upper (Hutchinson et al., 2017) body exercise. Importantly, the results of the current study provide support for the use of the $PRET_{max}$ in participants reliant on MWP.

This support for the use of $PRET_{max}$ comes from the finding of more positive affect during the $PRET_{max}$ compared to RAMP. This finding in participants reliant on MWP corroborates previous research using able-bodied participants performing recumbent cycle ergometry (Evans et al., 2014). The affect experienced during exercise could be a particularly important consideration when working with participants who have low habitual levels of physical activity, or who are unaccustomed to maximal exercise. This is based on evidence showing that affect during exercise predicted physical activity participation 6 to 12 months later in previously sedentary individuals (Williams et al., 2008). With the suggestion that more positive feelings during exercise can aid with adherence to a long-term exercise intervention, there are thus growing calls for the role of the affective response to receive greater consideration in exercise prescription guidelines (Ekkekakis, Parfitt, & Petruzzello, 2011; Williams, 2008). The current results would also support the consideration of affective response when selecting a maximal exercise test protocol.

This is the first study in this population group to challenge the traditional use of maximal incremental tests using fixed PO stages. The results also strengthen the case for obtaining a direct measurement of $\dot{V}O_{2peak}$ in contrast to predicting it from the $\dot{V}O_2$ at submaximal RPE, although only when maximal exercise testing is deemed safe and appropriate. Concerns over exacerbating the risk of shoulder injury, peripheral fatigue and autonomic dysfunction during maximal exercise has led to questions of whether maximal testing is appropriate in populations reliant on MWP (Totosy de Zepetnek, Au, Hol, Eng, & MacDonald, 2016). If though, as was the case in this study, maximal exercise is deemed safe then a direct measurement of $\dot{V}O_{2peak}$ should be made.

Previous studies have predicted $\dot{V}O_{2peak}$ from the submaximal $\dot{V}O_2$ during single-stage fixed PO (Totosy de Zepetnek et al., 2016), incremental fixed PO (Al-Rahamneh & Eston, 2011a; Al-Rahamneh et al., 2011; Goosey-Tolfrey et al., 2014) testing, and a submaximal PRET (Al-Rahamneh & Eston, 2011b). However mean difference (lower to upper limits of agreement) have been reported as 0 (-8 to 8) ml·kg^{-1}·min^{-1} (Al-Rahamneh & Eston, 2011a), 0.02 (-6.67 to 6.64) ml·kg^{-1}·min^{-1} (Totosy de Zepetnek et al., 2016), 0.4 (-5.3 to 6.1) ml·kg^{-1}·min^{-1} (Al-Rahamneh & Eston, 2011b) and 1 (-8 to 10) ml·kg^{-1}·min^{-1} (Al-Rahamneh et al., 2011) for various prediction models compared to 0.9 (-1.8 to 3.5) ml·kg^{-1}·min^{-1} as found in this study when comparing the $PRET_{max}$ and RAMP. These results show increased random error in the prediction models compared to the direct measurement from the $PRET_{max}$. Greater random error increases the possibility of a prediction that either under-, or over-, estimates $\dot{V}O_{2peak}$. These findings ultimately support the direct measurement of $\dot{V}O_{2peak}$ when possible, with the current study supporting the $PRET_{max}$ over a traditional RAMP.

In addition to the new knowledge around using RPE to prescribe the intensity during exercise testing for participants reliant on MWP, this study also adds support to the area of RPE-based exercise prescription for this population. The cost of equipment and technical expertise required for measuring $\dot{V}O_2$ and PO limit their use for informing exercise intensity away from a controlled laboratory setting. As such, individuals reliant on MWP have limited accessible methods for regulating exercise training intensity. It has been reported that there is currently insufficient evidence to support the regular use of subjective measures, such as RPE, to control intensity in adults with spinal cord injury (van der Scheer, Hutchinson, Paulson, Martin Ginis, & Goosey-Tolfrey, 2017). The present findings of comparable $\dot{V}O_{2peak}$ as well as RPE:$\dot{V}O_2$ and RPE:HR relationships between RAMP and $PRET_{max}$ protocols suggest that RPE may be used as a valid, cost effective and easily applicable means of prescribing exercise intensity in participants reliant on MWP. However as this study only investigates this using group-averaged single test relationships, further studies need to use a higher quality, estimation versus production study design to study this (van der Scheer et al., 2017).

A limitation of this study could be the sample size of 10 participants of a heterogeneous nature in terms of their mixed impairments, differing levels of cardiorespiratory fitness and habitual physical activity. Yet, despite the large inter-individual variation the findings showed that $PRET_{max}$ can be used to measure $\dot{V}O_{2peak}$ in persons reliant on MWP. Furthermore, while the participants had undertaken arm crank ergometry exercise before, several were unfamiliar with both the specific protocols ($PRET_{max}$ and RAMP), and indeed maximal exercise itself. This may have limited their ability to push themselves to achieve the intensity required (i.e. particularly for the $PRET_{max}$ final RPE 20 stage). This potentially manifested itself since the median RPE reported was 19 during this required RPE 20 stage of the test. The inability for these participants to apparently reach RPE 20 during the $PRET_{max}$, despite doing so in the RAMP, could serve to limit the $\dot{V}O_{2peak}$ values measured. Remarkably though, even with the difference in peak RPE reported between $PRET_{max}$ and RAMP, the $\dot{V}O_{2peak}$ values were shown to agree.

Conclusions

This is the first study to show that the $PRET_{max}$ can be used to reliably measure $\dot{V}O_{2peak}$ in participants reliant on MWP. Given the significantly more positive affect felt during the $PRET_{max}$ compared to RAMP, this study provides a compelling and convincing case for the use of the $PRET_{max}$ over RAMP in this population. The $PRET_{max}$ should be considered particularly when participants may be unaccustomed with maximal exercise and when the maximal exercise assessment is one of the first steps in prescribing a personalised exercise programme.

References

Al-Rahamneh, H., & Eston, R. G. (2011a). Prediction of peak oxygen consumption from the ratings of perceived exertion during a graded exercise test and ramp exercise test in able-bodied participants and paraplegic persons. *Archives of Physical Medicine & Rehabilitation*, *92*(2), 277–283.

Al-Rahamneh, H., & Eston, R. G. (2011b). The validity of predicting peak oxygen uptake from a perceptually guided graded exercise test during arm exercise in paraplegic individuals. *Spinal Cord*, *49*(3), 430–434.

Al-Rahamneh, H. Q., Faulkner, J. A., Byrne, C., & Eston, R. G. (2011). Prediction of peak oxygen uptake from ratings of perceived exertion during arm exercise in able-bodied and persons with poliomyelitis. *Spinal Cord*, *49*(1), 131–135.

Beltz, N. M., Gibson, A. L., Janot, J. M., Kravitz, L., Mermier, C. M., & Dalleck, L. C. (2016). Graded exercise testing protocols for the determination of VO_{2max}: historical perspectives, progress, and future considerations. *Journal of Sports Medicine*, *2016*, 1–12.

Bland, J. M., & Altman, D. G. (1996). Measurement error. *BMJ (Clinical Research Ed.)*, *313*(7059), 744.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*(2), 135–160.

Borg, G. A. (1998). *Borg's Perceived Exertion and Pain Scales*. Champaign, IL: Human Kinetics.

Chidnok, W., DiMenna, F. J., Bailey, S. J., Burnley, M., Wilkerson, D. P., Vanhatalo, A., & Jones, A. M. (2013). $\dot{V}O_{2max}$ is not altered by self-pacing during incremental exercise. *European Journal of Applied Physiology*, *113*(2), 529–539.

Dallmeijer, A. J., & van der Woude, L. H. (2001). Health related functional status in men with spinal cord injury: relationship with lesion level and endurance capacity. *Spinal Cord*, *39*(11), 577–583.

Ekkekakis, P., Parfitt, G., & Petruzzello, S. J. (2011). The pleasure and displeasure people feel when they exercise at different intensities: decennial update and progress towards a tripartite rationale for exercise intensity prescription. *Sports Medicine*, *41*(8), 641–671.

Evans, H., Parfitt, G., & Eston, R. (2014). Use of a perceptually-regulated test to measure maximal oxygen uptake is valid and feels better. *European Journal of Sport Science*, *14*(5), 452–458.

Goosey-Tolfrey, V., Paulson, T., Tolfrey, K., & Eston, R. (2014). Prediction of peak oxygen uptake from differentiated ratings of perceived exertion during wheelchair propulsion in trained wheelchair sportspersons. *European Journal of Applied Physiology*, *114*(6), 1251–1258.

Hamlyn-Williams, C. C., Freeman, P., & Parfitt, G. (2014). Acute affective responses to prescribed and self-selected exercise sessions in adolescent girls: an observational study. *BMC Sports Science, Medicine and Rehabilitation*, *6*, 35.

Hanson, N. J., Reid, C. R., Cornwell, K. M., Lee, T. L., & Scheadler, C. M. (2017). Pacing strategy during the final stage of a self-paced VO2max (SPV) test does not affect maximal oxygen uptake. *European Journal of Applied Physiology*, *117*(9), 1807–1815.

Hanson, N. J., Scheadler, C. M., Lee, T. L., Neuenfeldt, N. C., Michael, T. J., & Miller, M. G. (2016). Modality determines VO2max achieved in self-paced exercise tests: validation with the Bruce protocol. *European Journal of Applied Physiology*, *116*(7), 1313–1319.

Hardy, C. J., & Rejeski, W. J. (1989). Not what, but how one feels - the measurement of affect during exercise. *Journal of Sport & Exercise Psychology*, *11*(3), 304–317.

Hicks, A. L., Martin Ginis K.A., Pelletier, C. A., Ditor, D. S., Foulon, B., & Wolfe, D. L. (2011). The effects of exercise training on physical capacity, strength, body composition and functional performance among adults with spinal cord injury: a systematic review. *Spinal Cord*, *49*(11), 1103–1127.

Hill, D. W., Cureton, K. J., & Collins, M. A. (1989). Effect of time of day on perceived exertion at work rates above and below the ventilatory threshold. *Research Quarterly for Exercise and Sport*, *60*(2), 127–133.

Hjeltnes, N., & Jansen, T. (1990). Physical endurance capacity, functional status and medical complications in spinal cord injured subjects with long-standing lesions. *Paraplegia*, *28*(7), 428–432.

Hopkins, W. G. (2015). Spreadsheets for analysis of validity and reliability. Retrieved from http://www.sportsci.org/2015/ValidRely.htm

Hutchinson, M. J., Paulson, T. A. W., Eston, R., & Goosey-Tolfrey, V. L. (2017). Assessment of peak oxygen uptake during handcycling: test-retest reliability and comparison of a ramp-incremented and perceptually-regulated exercise test.

Plos ONE, *12*(7), e0181008.

Janssen, T. W., Dallmeijer, A. J., Veeger, D. J., & van der Woude, L. H. (2002). Normative values and determinants of physical capacity in individuals with spinal cord injury. *Journal of Rehabilitation Research and Development*, *39*(1), 29–39.

Leicht, C. A., Tolfrey, K., Lenton, J. P., Bishop, N. C., & Goosey-Tolfrey, V. L. (2013). The verification phase and reliability of physiological parameters in peak testing of elite wheelchair athletes. *European Journal of Applied Physiology*, *113*, 337–345.

Lim, W., Lambrick, D., Mauger, A. R., Woolley, B., & Faulkner, J. (2016). The effect of trial familiarisation on the validity and reproducibility of a field-based self-paced VO2max test. *Biology of Sport*, *33*(3), 269–275.

Middleton, J. W., Dayton, A., Walsh, J., Rutkowski, S. B., Leong, G., & Duong, S. (2012). Life expectancy after spinal cord injury: a 50-year study. *Spinal Cord*, *50*(11), 803–811.

Noakes, T. D. (2008). Testing for maximum oxygen consumption has produced a brainless model of human exercise performance. *British Journal of Sports Medicine*, *42*(7), 551–555.

Plichta, S. B., Kelvin, E. A., & Munro, B. H. (2013). *Statistical methods for health care research*. Philadelphia, PA: Lippincott Williams and Wilkins.

Rose, E. A., & Parfitt, G. (2007). A quantitative analysis and qualitative explanation of the individual differences in affective responses to prescribed and self-selected exercise intensities. *Journal of Sport & Exercise Psychology*, *29*(3), 281–309.

Savic, G., DeVivo, M. J., Frankel, H. L., Jamous, M. A., Soni, B. M., & Charlifue, S. (2017). Long-term survival after traumatic spinal cord injury: a 70-year British study. *Spinal Cord*, *55*(7), 651–658.

Straub, A. M., Midgley, A. W., Zavorsky, G. S., & Hillman, A. R. (2014). Ramp-incremented and RPE-clamped test protocols elicit similar VO2max values in trained cyclists. *European Journal of Applied Physiology*, *114*(8), 1581–1590.

Totosy de Zepetnek, J. O., Au, J. S., Hol, A. T., Eng, J. J., & MacDonald, M. J. (2016). Predicting peak oxygen uptake from submaximal exercise after spinal cord injury. *Applied Physiology, Nutrition, and Metabolism*, *41*(7), 775–781.

Valent, L., Dallmeijer, A., Houdijk, H., Talsma, E., & van der Woude, L. (2007). The effects of upper body exercise on the physical capacity of people with a spinal cord injury: a systematic review. *Clinical Rehabilitation*, *21*(4), 315–330.

van der Scheer, J. W., Hutchinson, M., Paulson, T., Martin Ginis, K. A., & Goosey-Tolfrey, V. L. (2017). Reliability and validity of subjective measures of aerobic intensity in adults with spinal cord injury: a systematic review. *PM&R*.

van Koppenhagen, C. F., Post, M., de Groot, S., van Leeuwen, C., van Asbeck, F., Stolwijk-Swuste, J., … Lindeman, E. (2014). Longitudinal relationship between wheelchair exercise capacity and life satisfaction in patients with spinal cord injury: A cohort study in the Netherlands. *The Journal of Spinal Cord Medicine*, *37*(3), 328–337.

Whipp, B. J., Davis, J. A., Torres, F., & Wasserman, K. (1981). A test to determine parameters of aerobic function during exercise. *Journal of Applied Physiology:*

Respiratory, Environmental and Exercise Physiology, *50*(1), 217–221.

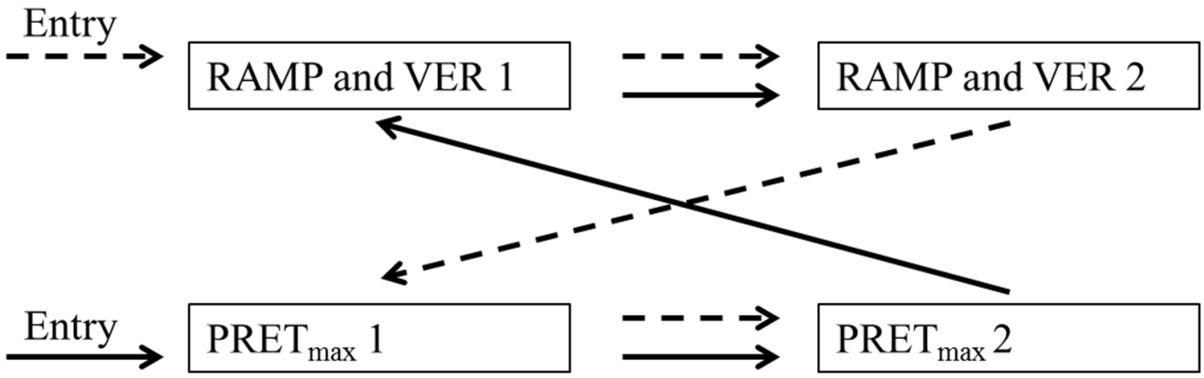
Williams, D. M. (2008). Exercise, affect, and adherence: an integrated model and a case for self-paced exercise. *Journal of Sport & Exercise Psychology*, *30*(5), 471–496.

Williams, D. M., Dunsiger, S., Ciccolo, J. T., Lewis, B. A., Albrecht, A. E., & Marcus, B. H. (2008). Acute affective response to a moderate-intensity exercise stimulus predicts physical activity participation 6 and 12 months later. *Psychology of Sport and Exercise*, *9*(3), 231–245.

Figure captions

Figure 1: Schematic of the experimental design used. Participants either performed ramp-incremented (RAMP) and verification (VER) trials in week 1 (dashed lines) followed by maximal perceptually-regulated exercise test ($PRET_{max}$) trials in week 2, or $PRET_{max}$ in week 1 (solid lines) followed by RAMP and VER in week 2.

Figure 2: Bland-Altman plots showing 95% LoA for a) absolute $\dot{V}O_{2peak}$, b) relative $\dot{V}O_{2peak}$, c) HR_{peak} and d) PO_{peak}. Mean difference between RAMP and $PRET_{max}$ trials is indicated by solid black line with upper and lower limits indicated by dotted lines.


Entry
RAMP and VER 1
RAMP and VER 2
Entry
$PRET_{max}$ 1
$PRET_{max}$ 2

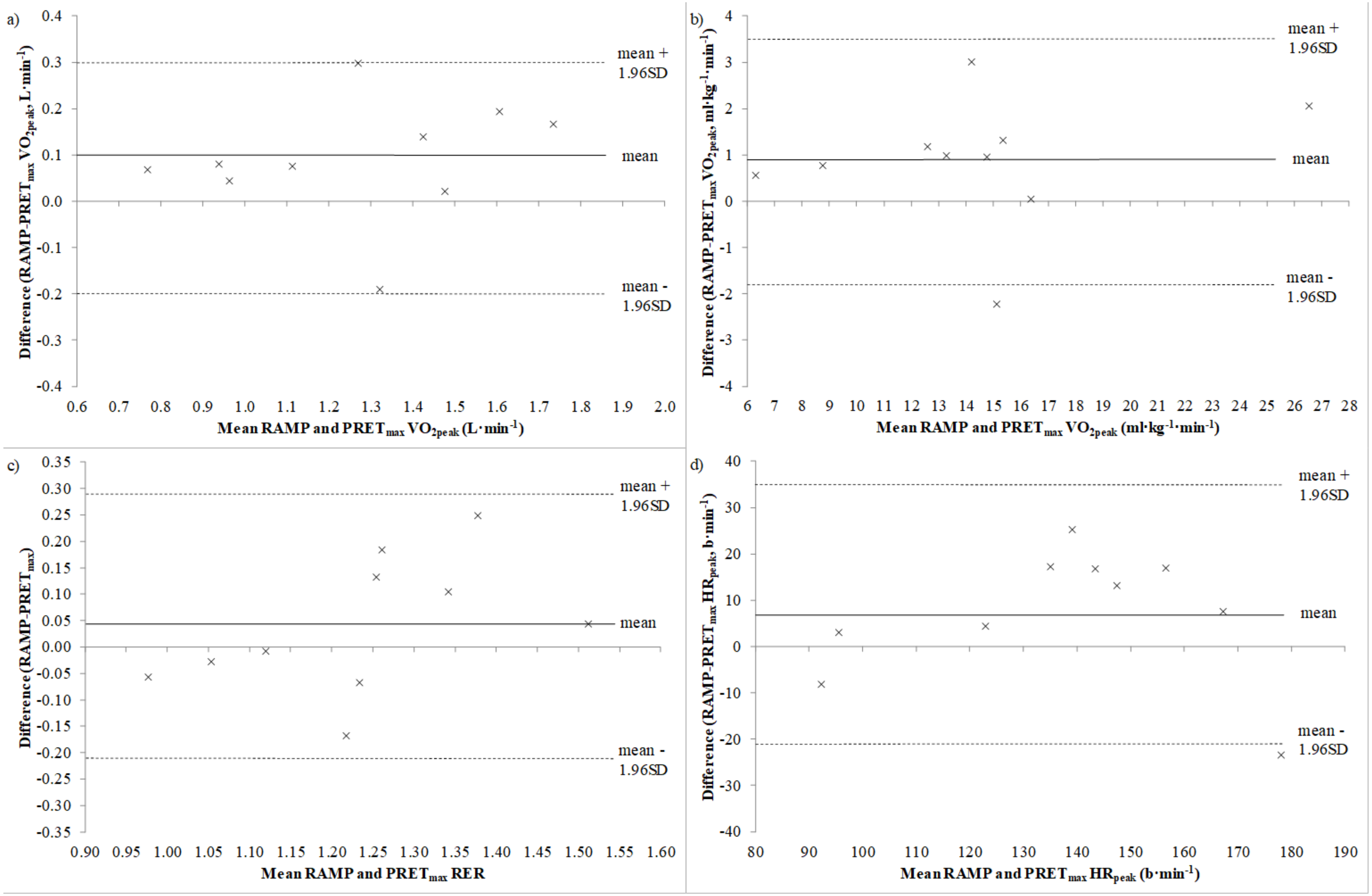

a)
Difference (RAMP-PRET$_{max}$ VO_{2peak}, L·min^{-1})
Mean RAMP and PRET$_{max}$ VO_{2peak} (L·min^{-1})
mean + 1.96SD
mean
mean - 1.96SD
b)
Difference (RAMP-PRET$_{max}$ VO_{2peak}, ml·kg^{-1}·min^{-1})
Mean RAMP and PRET$_{max}$ VO_{2peak} (ml·kg^{-1}·min^{-1})
mean + 1.96SD
mean
mean - 1.96SD
c)
Difference (RAMP-PRET$_{max}$)
Mean RAMP and PRET$_{max}$ RER
mean + 1.96SD
mean
mean - 1.96SD
d)
Difference (RAMP-PRET$_{max}$ HR_{peak}, b·min^{-1})
Mean RAMP and PRET$_{max}$ HR_{peak} (b·min^{-1})
mean + 1.96SD
mean
mean - 1.96SD

Table 1: Participant characteristics

Participant number	Gender	PAL (h·week^{-1})	Age (years)	Height (cm)	Body mass (kg)	Impairment	NLI	ASIA classification	TSI (years)
1	M	5	59	185	138.0	SCI	C4 (INC)	D	11
2	M	10	69	182	122.0	SCI	C5 (INC)	D	5
3	M	0	42	165	65.9	SCI	T5 (COMP)	A	17
4	M	5	75	178	92.4	SCI	T12 (INC)	D	10
5	M	0	49	170	61.6	SCI	L1 (INC)	D	11
6	M	0	57	193	90.6	SCI	L1 (INC)	D	18
7	M	5	63	183	87.5	MS			
8	M	10	50	178	91.1	MS			
9	M	6	60	190	107.0	MS			
10	F	4	48	145	85.0	Spina Bifida			
Mean		**4**	**57**	**177**	**94.1**				**12**
SD		**4**	**10**	**14**	**23.3**				**5**

ASIA: American Spinal Injury Association; COMP: complete; INC: incomplete; MS: multiple sclerosis; NLI: Neurological Level of Injury; PAL: Physical Activity Level; SCI: spinal cord injury; TSI: Time since injury.

Table 2: Test-retest reliability statistics for peak physiological variables obtained in RAMP and $PRET_{max}$ protocols.

	RAMP				$PRET_{max}$			
	CV (%)	ME	SDD	$ICC_{3,1}$	CV (%)	ME	SDD	$ICC_{3,1}$
$\dot{V}O_{2peak}$ ($L \cdot min^{-1}$)	4.6	0.12	0.16	0.95	5.4	0.13	0.18	0.93
$\dot{V}O_{2peak}$ ($ml \cdot kg^{-1} \cdot min^{-1}$)	5.9	1.70	2.36	0.96	8.1	2.20	3.04	0.92
HR_{peak} ($beats \cdot min^{-1}$)	3.8	11	15	0.95	3.7	8	12	0.97
PO_{peak} (W)	3.6	5	7	0.99	8.8	13	18	0.94
RER_{peak}	3.8	0.09	0.13	0.90	2.7	0.06	0.08	0.92

ME and SDD values are presented in the given unit of measurement for each variable. CV: Coefficient of Variation; HRpeak: peak heart rate; ICC: Intraclass correlation coefficient; ME: measurement error; POpeak: peak power output; PRETmax: maximal perceptually-regulated exercise test; RAMP: ramp-incremented max test; RERpeak: peak respiratory exchange ratio; SDD: smallest detectable difference; $\dot{V}O2peak$: peak oxygen uptake.

Table 3: Peak physiological and perceptual responses to RAMP and $PRET_{max}$.

	RAMP	$PRET_{max}$	Mean difference (95% CI)	*P*
$\dot{V}O_{2peak}$ (L·min^{-1})	1.3 ± 0.3	1.2 ± 0.3	0.1 (-0.0 - 0.2)	0.06
$\dot{V}O_{2peak}$ (ml·kg^{-1}·min^{-1})	14.8 ± 5.5	13.9 ± 5.2	0.9 (-0.1 - 1.8)	0.08
RER_{peak}	1.25 ± 0.20	1.22 ± 0.14	0.04 (-0.05 - 0.13)	0.37
HR_{peak} (beats·min^{-1})	141 ± 29	134 ± 29	7 (-3 - 18)	0.15
PO_{peak} (W)	81 ± 28	76 ± 34	6 (-3 - 14)	0.16
Duration (s)	674 ± 191	600 ± 0	74 (-63 - 210)	0.25
RPE_P	20 (19 - 20)*	19 (19 - 20)		0.03
RPE_C	20 (18 - 20)	19 (18 - 20)		>0.95
RPE_O	20 (18 - 20)	19 (19 - 20)		0.46
FS_{peak}	-3 (-4 - -1)*	0 (-2 - 1)		0.02
$FS_{average}$	2 (1 - 2)	2 (2 - 3)		0.21

Ratio data are presented as mean ± SD whilst ordinal data are presented as median (inter-quartile range). * = significantly different to PRETmax, $P < 0.05$. CI: confidence interval; FSpeak: peak Feeling Scale rating; FSaverage: average Feeling Scale rating; HRpeak: peak heart rate; POpeak: peak power output; PRETmax: maximal perceptually-regulated exercise test; RAMP: ramp-incremented exercise test; RER: peak respiratory exchange ratio; RPEc: central Rating of Perceived Exertion; RPEo: overall Rating of Perceived Exertion; RPEp: peripheral Rating of Perceived Exertion; V̇O2peak: peak oxygen uptake.

Table 4: Group-averaged correlations for differentiated RPE with objective markers of exercise intensity from the 2nd trial of each protocol

	RAMP	$PRET_{max}$
$RPE_P:\dot{V}O_2$	0.949 (0.769 - 0.990)	0.957 (0.640 - 0.996)
RPE_P:HR	0.967 (0.806 - 0.995)	0.971 (0.840 - 0.995)
RPE_P:PO	0.990 (0.947 - 0.998)*	0.970 (0.786 - 0.996)
$RPE_C: \dot{V}O_2$	0.956 (0.779 - 0.992)	0.954 (0.610 - 0.996)
RPE_C:HR	0.973 (0.788 - 0.997)	0.960 (0.641 - 0.996)
RPE_C:PO	0.991 (0.946 - 0.999)*	0.964 (0.610 - 0.997)
$RPE_O:\dot{V}O_2$	0.959 (0.676 - 0.996)	0.947 (0.458 - 0.996)
RPE_O:HR	0.969 (0.810 - 0.995)	0.959 (0.612 - 0.996)
RPE_O:PO	0.988 (0.932 - 0.998)*	0.965 (0.662 - 0.997)

Data are presented as mean (95% Confidence Interval). * = significant main effect of protocol for RPE:PO relationships, RAMP greater than PRETmax, $P < 0.05$. HR: heart rate; PO: power output; PRETmax: maximal perceptually-regulated exercise test; RAMP: ramp-incremented exercise test; RPEc: central Rating of Perceived Exertion; RPEo: overall Rating of Perceived Exertion; RPEp: peripheral Rating of Perceived Exertion; $\dot{V}O2$: oxygen uptake.

Table 5: Feeling Scale rating at submaximal RPE during RAMP and $PRET_{max}$

	RAMP	$PRET_{max}$
RPE 11	3.0 (2.9 - 4.0)	4.0 (3.5 - 4.5)
RPE 13	2.3 (1.0 - 3.0)	3.3 (2.9 - 3.6)*
RPE 15	0.8 (0.4 - 1.6)	2.0 (1.0 - 3.0)*
RPE 17	0.0 (-1.2 - 1.1)	1.5 (0.9 - 2.1)*

Data are presented as median (interquartile range). *: significantly greater than during RAMP, $P < 0.05$. PRETmax = maximal perceptually-regulated exercise test; RAMP = ramp-incremented exercise test; RPE = Rating of Perceived Exertion.