This is the author's version of a work that was accepted for publication in the following source:

Tan, J. L., T. Perera, J. L. McGinley, S. A. C. Yohanandan, P. Brown, and W. Thevathasan. 2018. Neurophysiological analysis of the clinical pull test. *Journal of Neurophysiology*. **120**(5): 2325-2333.

doi: 10.1152/jn.00789.2017

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source.

The final publication is available here

Copyright of this article belongs to: © 2018 the American Physiological Society

Neurophysiological analysis of the clinical pull test

Tan J.L[1,2], Perera T[1,3], McGinley J.L[3,4], Yohanandan S.A.C[3], Brown P[5], Thevathasan W[2,3,6]

1 Department of Medical Bionics, The University of Melbourne, Parkville, Victoria, Australia

2 Department of Neurology, The Royal Melbourne Hospital, Parkville, Victoria, Australia

3 The Bionics Institute, East Melbourne, Victoria, Australia

4 Department of Physiotherapy, The University of Melbourne, Parkville, Victoria, Australia

5 Medical Research Council Brain Network Dynamics Unit, and Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, United Kingdom.

6 Department of Medicine, The University of Melbourne, Parkville, Victoria, Australia

AUTHOR CONTRIBUTIONS

J.L.T., T.P., J.L.M., P.B., W.T. were involved in the conception and design of the experiments; J.L.T. and S.A.C.Y. performed experiments; J.L.T., T.P. and W.T, processed and analyzed the data; J.L.T., T.P., J.L.M, P.B., W.T. interpreted results of the experiments; all authors were involved in drafting or critical revision of the paper and approved the final version of manuscript.

Running Head

Neurophysiological analysis of the clinical pull test

Correspondence to: Dr Wesley Thevathasan. Department of Neurology, Royal Melbourne Hospital, Royal Parade, Parkville, Melbourne, 3050 Australia. Email;

wesley.thevathasan@mh.org.au

ABSTRACT

Postural reflexes are impaired in conditions such as Parkinson's disease, leading to difficulty walking and falls. In clinical practice, postural responses are assessed using the 'pull test', where an examiner tugs the pre-warned, standing patient backwards at the shoulders and grades the response. However, validity of the pull test is debated with issues including scaling and variability in administration and interpretation. It is unclear whether to assess the first trial or only subsequent repeated trials. The ecological relevance of a forewarned backwards challenge is also debated. We therefore developed an instrumented version of the pull test to characterize responses and clarify how the test should be performed and interpreted. In thirty-three healthy participants, 'pulls' were manually administered and pull force measured. Trunk and step responses were assessed with motion tracking. We probed for the StartReact phenomenon (where pre-prepared responses are released early by a startling stimulus) by delivering concurrent normal or 'startling' auditory stimuli. We found that the first pull triggers a different response, including a larger step size suggesting more destabilization. This is consistent with 'first trial effects', reported by platform translation studies, where movement execution appears confounded by startle reflex-like activity. Thus, first pull test trials have clinical relevance and should not be discarded as practice. Supportive of ecological relevance, responses to repeated pulls exhibited StartReact, as previously reported with a variety of other postural challenges including those delivered with unexpected timing and direction. Examiner pull force significantly affected the postural response particularly the size of stepping.

Keywords: Postural reflex, StartReact, Balance, Pull Test

NEW AND NOTEWORTHY

We characterized postural responses elicited by the clinical ‘pull test’ using instrumentation. The first pull triggers a different response, including a larger step size suggesting more destabilization. Thus, first trials likely have important clinical and ecological relevance and should not be discarded as practice. Responses to repeated pulls can be accelerated with a startling stimulus, as reported with a variety of other challenges. Examiner pull force was a significant factor influencing the postural response.

INTRODUCTION:

Postural ‘reflexes’ are crucial to the maintenance of upright stance (Currie and Carlsen 1985; Eaton and Emberley 1991; Shemmell 2015). Impairment of these reflexes in disorders such as Parkinson’s disease results in postural instability and causes reduced walking confidence, falls and even inability to stand (Bloem et al. 2004; Hely et al. 2005). In clinical practice, postural responses are routinely assessed using the ‘pull test’. In the pull test, an examiner tugs the standing patient backwards at the shoulders and grades the response (Fahn S, Elton RL. UPDRS Development Committee 1987). Patients typically respond by flexing at the trunk and sometimes by taking a corrective step (Visser et al. 2003; Hunt and Sethi 2006). However, pull test scores correlate poorly with important clinical endpoints such as falls and poor validity of the pull test is cited as a factor confounding the detection of novel treatments (Bloem et al. 1998; Thevathasan et al. 2011; Morita et al. 2014). Such issues could simply reflect the limited scaling of the test (score/4). Additionally, inter and intra rater reliability are often cited as confounds (Munhoz et al. 2004; Hunt and Sethi 2006). However, it is unclear whether variabilities in test administration, such as pull force, influence the response. Interpretation of the test is also controversial. For example, it is debated which trial to assess. Guidelines produced by the International Movement Disorders Society suggest that the patient should be fore-warned about the impending challenge, with an initial practice trial before a second trial is formally assessed (Goetz et al. 2008). Others claim that an unexpected pull, performed only once, is most clinically meaningful (Bloem et al. 2001; Visser et al. 2003). In some studies, the average response from repeated pull test trials has been measured (Bloem et al. 2001; Nanhoe-Mahabier et al. 2012). Whether these different techniques yield different motor responses is uncertain. For example, in contrast to repeated trials, initial perturbations could be more ‘startling’ and the motor response less pre-prepared (Allum et al. 2011). This also raises the question of ecological relevance. An expected backwards

perturbation as occurs in the pull test may occur relatively infrequently in daily life, such as a passenger subject to anticipated acceleration of a train. Do such responses differ from those triggered by challenges that occur unexpectedly and in any direction (Carpenter et al. 2004; Dimitrova et al. 2004; Jacobs et al. 2005)?

Some of these questions have been addressed in laboratory studies, using a different perturbation – platform translation beneath the feet (Campbell 2012; Campbell et al. 2012, 2013; Nonnekes et al. 2013; Ravichandran et al. 2013). These studies suggest that first and subsequent postural challenges evoke different responses. “First trial postural responses” have many features that suggest superimposition of the generalized startle reflex, including muscle co-contraction, forward flexion of the trunk into a crouching posture and subsequent habituation (Oude Nijhuis et al. 2009, 2010; Allum et al. 2011). This ‘first trial effect’ may actually be maladaptive, at least for stance preservation, being associated with increased deviations in centre of mass and propensity to falls (Horak and Nashner 1986; Oude Nijhuis et al. 2009). If such findings also apply to the pull test, this would be of great relevance to clinicians, suggesting that the very first pull test trial is crucial to assess and should not be discarded as practice. This may be particularly true in patients with Parkinson’s disease, who are reported to have abnormalities integrating startle responses into movement (Nieuwenhuijzen et al. 2006). Platform translation studies also suggest that postural responses, averaged over repeated trials, exhibit pre-preparation but keep in reserve latent potential to be accelerated (Campbell et al. 2012; Nonnekes et al. 2013). Pre-preparation was assessed in these studies using the StartReact probe – where a concurrent ‘startling’ stimulus, such as a loud sound, can speed responses that are pre-prepared (Valls-Solé et al. 1995; Valls-Sole et al. 1999; Campbell 2012; Nonnekes et al. 2013). Platform translation studies report that StartReact is present with backwards and sideways challenges and regardless of whether the perturbation is expected in timing or direction (Campbell et al. 2012; Nonnekes

et al. 2013). If repeated pull test trials also exhibit StartReact, this could support that pull test results reflect the integrity of pathways utilized by a broader range of responses than only those elicited by a forewarned backwards challenge.

However, whether findings from platform translation studies apply to the pull test is uncertain. Platform translation involves intense 'bottom up' perturbations, generating an initial lower limb response as occurs when slipping on a wet floor (Horak et al. 1997). In contrast, the pull test employs a 'top down' perturbation, with initial displacement and response of the trunk, as may occur when bumped in a crowd – and this generates a different pattern of motor recruitment (Govender et al. 2015; Azevedo et al. 2016; Colebatch et al. 2016; Di Giulio et al. 2016).

We therefore developed an instrumented version of the pull test to better characterize the nature of elicited postural responses and clarify how the test should be performed and interpreted. Like the clinical test, the perturbation was delivered manually by an examiner but with measurement of pull force. Both the trunk and step responses were assessed with motion tracking – akin to visual assessment by a clinician. Three key questions were addressed. First, whether responses to the first pull test trial differ from subsequent repeated trials. Secondly, whether averaged responses to repeated pulls exhibit StartReact. Third, whether variabilities in baseline subject characteristics such as height and weight, or in the examiner such as pull force, affect results.

PARTICIPANTS AND METHODS

Participants:

Thirty-three healthy young adults (age 28.0 ± 4.1 years; height 1.72 ± 0.1 m; weight 68.8 ± 14.0 kg; 20 males) without known hearing, neurological or musculoskeletal disorders were recruited as a sample of convenience. Repeated trial data were captured for all participants. In

eighteen participants, first trial data were also captured, and pulls were of sufficient force to always generate both a trunk and step response. In the remaining fifteen participants, pulls were of lesser force and elicited only a trunk response. Local ethics committee approval was obtained, and participants gave written informed consent.

Experiments:

The instrumented pull test was performed similarly to the clinical pull test (Fahn S, Elton RL. UPDRS Development Committee 1987). Participants stood in bare feet and focused on a picture 1.5 m ahead at eye level, wearing a customized trunk harness attached to a load cell (Omegadyne LCM201-100N, Ohio, USA). A warning cue was not provided. The assessor manually generated a backwards pull via the load cell held perpendicular to the shoulder level of the participant (Fig. 1). Thirty-five trials were presented serially, with an auditory stimulus (40 ms duration, 1000 Hz) delivered within 30 ms of each pull. The auditory stimulus was either 90 dB (normal) or 116 dB (loud). This loud stimulus has been demonstrated as sufficient to trigger StartReact and a startle reflex (Thevathasan et al. 2011). The first five trials involved normal sounds, followed by 20 normal and 10 loud trials randomly intermixed. Inter-trial intervals (10 - 15 s) were variable. Participants had a short rest after each block of 10 pulls, or as requested.

Auditory stimuli were delivered through a custom hardware audio interface. Auditory tones were delivered binaurally through headphones (Audio Technica, ATH-ES7, Tokyo, Japan). Sound pressure levels were calibrated in a sound-proof room with a modular precision sound analyzer (Observer 2260, Bruel and Kjaer, Naerum, Denmark) via an artificial ear and headphone adaptor. Tones were triggered when the force of each pull reached a maximum (Fig. 2). Triggering was processed within the hardware via an embedded microprocessor, resulting in a delay of 21 ± 6 ms between the onset of the pull and auditory stimulus

delivery. This delay is well within the time window where a loud auditory stimulus can trigger StartReact (Valls-Solé 2004; Kumru and Valls-Solé 2006; Castellote et al. 2013).

Responses in both tasks were measured with electromagnetic three-dimensional motion tracking using type-800 miniature sensors (Ascension TrakStar, Vermont, USA). Sensors were attached at the sternal notch (at the level of the second and third thoracic vertebra), and on the right and left ankle malleolus. Each sensor, sampled at 250 Hz, measured triaxial displacement in millimeter units as well as pitch, roll, and yaw in degrees. The sensors were referenced to the origin of the transmitter. Data were acquired using custom software. Previous work has shown this system has a sensitivity of 0.45 mm and 0.02° with measurements accurate to ± 0.4 mm and ± 0.05° (Perera et al. 2016).

Tasks were administered in a quiet room with distractions minimized. The order of tasks was counterbalanced. Participants were blinded to experiment hypotheses. One researcher (J.L.T.) conducted all assessments and continuously monitored participants to prevent falls.

Parameters

Data analysis was automated using a script written in MATLAB (MathWorks Inc., Massachusetts, USA). Motion tracking data were high-pass filtered with a 0.05 Hz cut-off frequency. Trunk displacement data were differentiated to determine velocity and then acceleration.

The truncal response is the first strategy used to maintain upright posture in the pull test and is elicited in every trial (Di Giulio et al. 2016). Postural reaction time was defined as the difference between the onset of trunk displacement and the turning point of the trunk velocity curve (when the trunk started to decelerate) (Fig. 2 and 3). Notably, the imperative was defined as the onset of trunk displacement (three standard deviations above the 1 s pre-stimulus baseline), rather than any threshold of pull force - as this removed the confound of

variability due to slack in the harness or body. Magnitude of the postural response was defined as the peak deceleration of the trunk and is reported in units of millimeters per second per second ($mm.s^{-2}$).

When the truncal response is insufficient to maintain balance, the step response is generated (Pai et al. 1998). Stepping can also occur early, well within stability limits and before balance is fully perturbed (Maki and McIlroy 1997). Stepping was defined as the foot moving past the stance foot in the backward direction, excluding movement in any other direction. Step reaction time was calculated as the difference between the onset of truncal displacement to the initial movement of the stepping limb (4 standard deviations above baseline). Response magnitude of stepping was determined by total displacement of the feet in millimeters (mm), from initial foot lift off to contact of the stepping limb arresting backward retropulsion. Analysis excluded steps less than 50 mm, as the change in base of support is considered negligible (McVey et al. 2013).

Reaction times and response magnitudes were computed for every individual trial including the first trial. For repeated trials, the first five trials were discarded to avoid first trial effects which have been shown to habituate over several initial trials (Nanhoe-Mahabier et al. 2012). Repeated trial measures reflect averages over the subsequent thirty trials. Additionally, trials were rejected if there was an anticipatory truncal movement (a forward trunk displacement immediately prior to the auditory stimulus that exceeds three standard deviations above the 1s pre-stimulus baseline). In one participant, anticipatory truncal movement was detected in most trials, so this dataset was excluded.

Peak pull force and rate of force development were calculated from the load cell and reported in units of Newtons (N) and Newtons per second (N/s) respectively. The peak pull force

indicates the instantaneous maximum force delivered, whereas the force rate is the slope of the force versus time curve indicating how rapidly the force was generated.

Data Analysis

A Kolmogorov–Smirnov Test demonstrated that all measures did not differ from a normal distribution.

Due to the number of contributing factors that could influence trunk and stepping reaction time and magnitude in the postural reflex task, a linear mixed models (LMM) analysis was conducted according to methods previously described (Boisgontier et al. 2017). LMM offers several advantages over ANOVA in the analysis of the postural reflex task by accounting for both nested (multiple observations in a single participant in one condition) and crossed (participant observed in multiple conditions) factors, controlling for increased probability of Type 1 error (Boisgontier and Cheval 2016). At the participant level, it accounts for height and weight, and sampling variability of peak force and force rate at the trial level. LMMs prevent information loss by considering all trials individually, rather than averaging data across multiple trials. Results can subsequently be generalized across the population and conditions tested (Boisgontier et al. 2017).

To determine effects of auditory stimulus and pull force variables on reaction time and response magnitude in the postural reflex task, LMM analysis was conducted using the following equation:

$$\mathrm{Y}_{ij} = \left(\beta_0 + \theta_{0j}\right) + \beta_1 \mathrm{TrialType}_{ij} + \beta_2 \mathrm{Weight}_j + \beta_3 \mathrm{Height}_j + \beta_4 \mathrm{PeakForce}_{ij} + \beta_5 \mathrm{ForceRate}_{ij} + \epsilon_{ij}$$

Where $\mathrm{Y}_{ij}$is the participant's reaction time or response magnitude for trial i, β_{0-5} are the fixed effect coefficients, θ_{0j} is the random effect for participant j (random intercept), and ϵ_{ij} is the error term.

This model was built using SPSS (IBM, Chicago, IL) version 22. The following factors were included: trial type, weight, height, peak force, and force rate. At the participant level, height and weight were included to account for the lever-arm mechanism of the pull and the impact inertia of the trunk (Delitto et al. 1989; Oliveira et al. 2011). To determine if the first trial was different from subsequent trials, the trial factor is coded as 0 for the first trial (90 dB), 1 for repeated trials at 90 dB, and 2 for repeated trials at 116 dB. Four LMMs relating to: trunk reaction time, trunk response magnitude, stepping reaction time, and step response were used to analyze the data. Variance inflation factors for each predictor in all models fell below 10 and multicollinearity was considered absent (Hair et al. 1995). The variance components covariance structure was selected for the LMMs.

For each explanatory variable, an estimate of the effect, p-value, and 95% confidence intervals are specified. Where the explanatory variable is continuous, the estimate of the effect is based on a regression coefficient (which gives the predicted increase in outcome for a one-point increase in the variable). For categorical explanatory variables, post-hoc pairwise comparisons were performed to determine the differences between trial types and were corrected for multiple comparisons (Benjamini and Hochberg 1995). Level of significance was $\alpha = 0.05$.

RESULTS

A summary of results from the LMM analysis is found in the tables.

First trial responses

Trunk reaction times from first trials did not differ from repeated trials with normal stimuli [$p = 0.692$] and repeated trials with loud stimuli [$p = 0.692$]. Trunk response magnitudes from

first trials did not differ from repeated trials with normal stimuli [p = 0.497] and repeated trials with loud stimuli [p = 0.120].

Step reaction times from first trials were slower than repeated trials with normal stimuli [mean difference 36.9 ms, p = 0.009] and repeated trials with loud stimuli [mean difference 46.1 ms, p = 0.002]. Step response magnitudes from first trials were larger than repeated trials with normal stimuli [mean difference 60 mm, p = 0.002] and repeated trials with loud stimuli [mean difference 53 mm, p = 0.005].

StartReact in repeated trials

StartReact was present in the trunk response from repeated trials. That is, trunk reaction times in repeated trials were faster with loud compared with normal stimuli [mean difference 10.2 ms, p = 0.002]. Trunk response magnitude in repeated trials were also larger with loud than normal stimuli [mean difference 588 $mm.s^{-2}$, p < 0.001]. However, there was no difference in repeated trials with normal and loud stimuli for step reaction time [p = 0.072] and step response magnitude [p = 0.315].

Impact of examiner and baseline participant variables

Increased peak pull force was associated with slower trunk reaction times [b = 0.36, p < 0.001] and larger step response magnitude [b = 1.02, p < 0.001]. Increased participant weight was associated with slower step reaction time [b = 2.37, p = 0.008]. Otherwise, participant weight and height did not influence results.

DISCUSSION

We sought to characterize the nature of postural responses generated by the pull test, to clarify how the test should be performed and interpreted. First pull test trials were significantly different to subsequent repeated trials, demonstrating a slower and larger step response. Despite the relative ‘surprise’ of first trial perturbations (all with normal stimuli),

the trunk response was no faster or larger than in repeated trials (with either normal or loud stimuli). In repeated trials, the trunk demonstrated StartReact – that is, loud stimuli were associated with faster responses. Increased peak pull force increased trunk response latency and increased the size of the step response. Increased subject weight increased step response latency. Otherwise, subject weight and height did not influence results.

Before further discussion, potential confounds need to be considered. It is important to note that in this study, we employed motion tracking as the assessment tool – which detects net movement (akin to a clinician) rather than when muscle recruitment first begins - as is available with electromyography (EMG). Importantly, for first trials, we only employed normal stimuli and not also loud stimuli. Thus, we cannot directly assess whether first trials exhibited StartReact. One question is whether intensity effects (Woodworth 1938; Kohfeld 1969) contributed to our finding of StartReact in repeated trials. However, previous studies (Carlsen et al. 2007; Delval et al. 2012) that employed similar stimulus intensities report that the substantial reaction time benefit of StartReact is inconsistent with the modest and gradual reduction in reaction times seen with increasing stimulus intensities. Another potential confound is that of intersensory facilitation (Hershenson 1962). In this study, we excluded intersensory facilitation by employing a control sound of 90 dB in normal trials in addition to the potentially ‘startling’116 dB in loud trials (Thevathasan et al. 2011).

Characterization of postural responses to the pull test

First trial effects

We found that first pull test trials (with normal stimuli) differed significantly from repeated trials (with normal and loud stimuli). Whilst first trial trunk responses did not differ in latency or magnitude compared with repeated trials with normal and loud stimuli (see further discussion below), the step response was slower and the step magnitude bigger. We interpret

this larger step as indicating that first pull test perturbations had a more destabilizing impact on stance. This is supported by our observation that greater peak pull force also generated a larger step response. This also corroborates findings that first trials provoked by platform translation result in greater displacement in centre of mass and increased falls (Horak and Nashner 1986; Oude Nijhuis et al. 2009). Why would first trial responses be performed worse than subsequent repeated trials? First trials are performed slower even for a non-postural, non-startling task of ankle dorsiflexion and this may reflect the lesser opportunity for motor preparation compared with after practice (Sutter et al. 2016). This could explain why we found the step response in first trials to be slower (see further discussion below). However, a more specific issue for first *postural* challenges are findings reminiscent of a confounding startle reflex including excessive muscle co-contraction, crouching body response and subsequent habituation (Oude Nijhuis et al. 2009, 2010). The exact nature of the first trial effect is debated, and the pattern of muscle recruitment may not simply be explained by summation of the startle reflex and postural response (Oude Nijhuis et al. 2010). Nevertheless, like the startle reflex, the impact of first trial postural responses to stiffen and crouch the body may have provided evolutionary advantage to our ancestors. However, for the maintenance of upright stance, the first trial response appears to be maladaptive (Allum et al. 2011). In this study, we did not seek the full array of features characterizing the first trial postural response. However, discrimination of first trial from repeated postural responses can be difficult, for example activation of sternocleidomastoid and masseter occurs in both (Oude Nijhuis et al. 2010; Campbell et al. 2013). Regardless, for the pull test we have found that first trial responses are different to repeated trial responses. This supports that the first pull test trial is worth assessing as a separate endpoint – and could be argued to have greater ecological relevance (see below).

StartReact and motor preparation

For repeated trials, we found that postural responses of the trunk exhibited StartReact, supporting the existence of pre-preparation (Valls-Solé et al. 1995; Valls-Sole et al. 1999). Only a weak trend suggested StartReact in the step response – which may reflect the secondary importance of stepping; being required only when the earlier trunk response is insufficient (Pai et al. 1998). That StartReact is present in the postural response of the trunk in repeated pull test trials is perhaps unsurprising, given these trials benefitted from practice and foreknowledge of the required response. These findings corroborate platform translation studies which report that repeated postural responses to both forward and sideways perturbations exhibit StartReact (Campbell 2012; Nonnekes et al. 2013). This suggests that repeated postural responses keep in reserve capacity to be released even quicker when there is a concurrent and suitably arresting stimulus. At least one study found that provoking this latent StartReact effect in postural responses to be advantageous, improving balance preservation (Nonnekes et al. 2013). However, should not first postural trials also trigger StartReact (even in the absence of a loud stimulus) given their importance and propensity to generate a startle-like response? Indeed, one study found that a first postural challenge was itself a sufficiently arresting stimulus to provoke StartReact in wrist extension (Campbell et al. 2013). However, whether the postural response itself benefits from StartReact in first trials has not been explored (Campbell et al. 2013). Our results, on superficial review, suggest that this may not occur – as we found that first trial trunk reaction times were not different from repeated trials with normal stimuli. However, we did not compare postural responses to normal and loud sounds *in first trials*. Moreover, a recent study found that for a non-postural non-startling ankle dorsiflexion task, first trials were actually performed *slower* than subsequent repeated trials (Sutter et al. 2016). Loud startling stimuli sped these first dorsiflexion trials so that they had a similar latency to repeated trials with normal stimuli (Sutter et al. 2016). Taken together, these findings raise the possibility that StartReact could

actually have benefitted our first pull test trunk responses but sufficed only to bring latencies in line with repeated trials with normal stimuli. This would also explain why the step response, which did not clearly benefit from StartReact in our study, remained slower in first trials compared with repeated trials. This is a hypothesis that could be addressed in future research.

Impact of examiner and baseline participant variables

Peak pull force varied sufficiently to be a factor affecting the latency of the trunk response and the size of stepping. The slower trunk response with increased peak pull force may reflect the greater recruitment of trunk muscles required before sufficient counteracting acceleration (deceleration) could be generated (our definition of response onset) (Cresswell et al. 1994). That increased peak pull force was associated with larger step size likely reflects the greater destabilization produced by a more forceful perturbation and thus the magnitude of the compensatory step required (Pai et al. 1998). Increased participant weight resulted in slower step reaction times. A relationship between increased body mass and slower reaction time has previously been reported (Skurvydas et al. 2009). This may be, at least partly, explained by Newton's second law (Force = Mass x Acceleration). That is, increased recruitment of leg muscles may be required to generate sufficient force to produce step acceleration, when mass is greater (Skurvydas et al. 2009). Otherwise, weight and height did not influence pull test performance although this may reflect limited variance of these parameters in the participants.

Clinical relevance

We tested a cohort of young healthy controls and whether these findings are applicable to older patients with Parkinson's disease is worthy of future investigation. For example, StartReact is reported to be delayed with ageing (Tresch et al. 2014) and is reported to be absent in some patients with Parkinson's disease (Thevathasan et al. 2011). Regardless, our

results are likely to have interest for clinicians who perform the pull test. Our findings support that the first trial response is important to capture as an endpoint in its own right rather than to be ignored in the primary assessment of the second trial (as suggested in the Movement Disorders Society Unified Parkinson's Disease Rating Scale) or averaged out in analysis of repeated trials (Visser et al. 2003; Goetz et al. 2008; Oude Nijhuis et al. 2010; Nonnekes et al. 2015). Interestingly, previous studies have suggested that the first trial may correlate best with other clinical measures of balance impairment such as falls (Visser et al. 2003; Oude Nijhuis et al. 2010; Nanhoe-Mahabier et al. 2012; Nonnekes et al. 2015). It is encouraging and supportive of ecological relevance, that postural responses to the pull test exhibit similar characteristics (at least in terms of first trial responses and StartReact in repeated trials) to those generated in the laboratory from bottom up perturbations where the timing and directions of perturbations are unknown (Campbell 2012; Nonnekes et al. 2013). On the other hand, it seems clear that examiner performance, namely peak pull force, has a substantial impact on pull test results. Clinicians do not have the benefit of a pull force meter and a mixed linear model to adjust for such confounds. Whilst the pull test remains a very useful clinical test, these findings help explain why reliability has been an issue and supports the call to develop more objective measures or biomarkers of postural instability (Perera et al. 2018).

Our results may have bearing on rehabilitation strategies that have been employed in patients with Parkinson's disease whereby patients are subject to repeated pulls to enhance the practiced response (Dijkstra et al. 2015; Peterson et al. 2016). It is unclear if this approach would benefit the first, and arguably most important, postural response given the propensity for first trial effects to confound movement execution. These first trial responses may be exaggerated in patients with Parkinson's disease, who have delays in habituation compared to controls (Nanhoe-Mahabier et al. 2012). It has been speculated that exaggerated first trial

responses in Parkinson's disease patients may arise as a consequence of fear of falling (Grillon et al. 1991; Davis et al. 1993; Adkin et al. 2003, 2008; Franchignoni et al. 2005). Furthermore, one study (Nieuwenhuijzen et al. 2006) observed that patients with Parkinson's disease are less able to integrate the startle response into phases of gait. If diminished integration of the startle reflex can also involve first postural responses – then this could conceivably be a factor in the risk of falls in patients with Parkinson's disease. If so, then it is worth noting that therapies exist to suppress startle and therefore falls in patients with hyperekplexia (exaggerated startle) (McAbee 2015).

The Instrumented Pull Test as a Potential Assessment Tool

This study highlights the capabilities of instrumentation of the clinical pull test with responses assessed with motion tracking. As in clinical practice, the perturbation was delivered manually by an examiner. To deliver the pull, we employed a rope attached to a harness with a force gauge to record the force of each pull. The recording of pull force appears important, as this was a significant cofactor affecting results. The use of an external motion tracker to measure responses was akin to visualization of movement used by clinicians. Of crucial importance was the method of analyzing motion tracking data with respect to the onset of trunk displacement rather than the pull itself. This decision to time lock to the onset of trunk displacement allowed us to exclude several confounds including the variable time taken to tension the harness and rope by the assessor and the variable stiffening of the body between trials.

Recently, more precisely calibrated truncal perturbations have been attempted in the laboratory with motors and pendulums (Azevedo et al. 2016; Di Giulio et al. 2016). In contrast, motion tracking is a relatively simple technique (albeit still requiring specialized equipment) and this has been previously employed to assess the 'push and release test' (an

alternative to the more widely used ‘pull test’)(Jacobs et al. 2006; Smith et al. 2016). The instrumented pull test reported here could therefore be a more accessible alternative to assess patients with Parkinson’s disease for clinical research. However, we note that the instrumented pull test described here would only be able to finely grade responses in patients who can still maintain stance without falling – that is, patients with Parkinson’s disease (up to Hoehn and Yahr stage 3/5) who exhibit up to grade 1 postural instability according to the Unified Parkinson’s disease Rating Scale (Fahn S, Elton RL. UPDRS Development Committee 1987).

ACKNOWLEDGEMENTS

Dr. Sue Finch (Statistical Consulting Centre and Melbourne Statistical Consulting Platform, University of Melbourne) provided statistical support for this study.

GRANTS

This work was supported by funding through the National Health and Medical Research Council (1066565), the Victorian Lions Foundation, and The Victorian Government’s Operational Infrastructure Support Program.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

REFERENCES

Adkin AL, **Campbell AD**, **Chua R**, **Carpenter MG**. The influence of postural threat on the cortical response to unpredictable and predictable postural perturbations. *Neurosci Lett* 435: 120–125, 2008.

Adkin AL, **Frank JS**, **Jog MS**. Fear of falling and postural control in Parkinson's disease. *Mov Disord* 18: 496–502, 2003.

Allum JHJ, **Tang K-S**, **Carpenter MG**, **Oude Nijhuis LB**, **Bloem BR**. Review of first trial responses in balance control: Influence of vestibular loss and Parkinson's disease. *Hum Mov Sci* 30: 279–295, 2011.

Azevedo AK e C de, **Claudino R**, **Conceição JS**, **Swarowsky A**, **Santos MJ dos**. Anticipatory and Compensatory Postural Adjustments in Response to External Lateral Shoulder Perturbations in Subjects with Parkinson's Disease. *PLOS ONE* 11: e0155012, 2016.

Benjamini Y, **Hochberg Y**. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* 57: 289–300, 1995.

Bloem BR, **Beckley DJ**, **van Hilten BJ**, **Roos RA**. Clinimetrics of postural instability in Parkinson's disease. *J Neurol* 245: 669–673, 1998.

Bloem BR, **Grimbergen YAM**, **Cramer M**, **Willemsen M**, **Zwinderman AH**. Prospective assessment of falls in Parkinson's disease. *J Neurol* 248: 950–958, 2001.

Bloem BR, **Hausdorff JM**, **Visser JE**, **Giladi N**. Falls and freezing of gait in Parkinson's disease: A review of two interconnected, episodic phenomena. *Mov Disord* 19: 871–884, 2004.

Boisgontier MP, **Cheval B**. The anova to mixed model transition. *Neurosci Biobehav Rev* 68: 1004–1005, 2016.

Boisgontier MP, **Cheval B**, **Chalavi S**, **van Ruitenbeek P**, **Leunissen I**, **Levin O**, **Nieuwboer A**, **Swinnen SP**. Individual differences in brainstem and basal ganglia structure predict postural control and balance loss in young and older adults. *Neurobiol Aging* 50: 47–59, 2017.

Campbell AD. Insights into human dynamic balance control : postural response initiation explored through classical conditioning and startle [Online]. http://circle.ubc.ca/handle/2429/43307 [19 Jan. 2015].

Campbell AD, **Chua R**, **Inglis JT**, **Carpenter MG**. Startle induces early initiation of classically conditioned postural responses. *J Neurophysiol* 108: 2946–2956, 2012.

Campbell AD, **Squair JW**, **Chua R**, **Inglis JT**, **Carpenter MG**. First trial and StartReact effects induced by balance perturbations to upright stance. *J Neurophysiol* 110: 2236–2245, 2013.

Carlsen AN, **Dakin CJ**, **Chua R**, **Franks IM**. Startle produces early response latencies that are distinct from stimulus intensity effects. *Exp Brain Res* 176: 199–205, 2007.

Carpenter M, **Allum J**, **Honegger F**, **Adkin A**, **Bloem B**. Postural abnormalities to multidirectional stance perturbations in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 75: 1245–1254, 2004.

Castellote JM, **Van den Berg MEL**, **Valls-Sole J**. The StartReact Effect on Self-Initiated Movements. *BioMed Res Int* 2013, 2013.

Colebatch JG, **Govender S**, **Dennis DL**. Postural responses to anterior and posterior perturbations applied to the upper trunk of standing human subjects. *Exp Brain Res* 234: 367–376, 2016.

Cresswell AG, **Oddsson L**, **Thorstensson A**. The influence of sudden perturbations on trunk muscle activity and intra-abdominal pressure while standing. *Exp Brain Res* 98, 1994.

Currie S, **Carlsen RC**. A rapid startle response in larval lampreys. *Brain Res* 358: 367–371, 1985.

Davis M, **Falls WA**, **Campeau S**, **Kim M**. Fear-potentiated startle: a neural and pharmacological analysis. *Behav Brain Res* 58: 175–198, 1993.

Delitto A, **Crandell CE**, **Rose SJ**. Peak Torque-to-Body Weight Ratios in the Trunk: A Critical Analysis. *Phys Ther* 69: 138–143, 1989.

Delval A, **Dujardin K**, **Tard C**, **Devanne H**, **Willart S**, **Bourriez J-L**, **Derambure P**, **Defebvre L**. Anticipatory postural adjustments during step initiation: elicitation by auditory stimulation of differing intensities. *Neuroscience* 219: 166–174, 2012.

Di Giulio I, **St George RJ**, **Kalliolia E**, **Peters AL**, **Limousin P**, **Day BL**. Maintaining balance against force perturbations: impaired mechanisms unresponsive to levodopa in Parkinson's disease. *J. Neurophysiol.* (April 20, 2016). doi: 10.1152/jn.00996.2015.

Dijkstra BW, **Horak FB**, **Kamsma YPT**, **Peterson DS**. Older adults can improve compensatory stepping with repeated postural perturbations. *Front Aging Neurosci* 7: 201, 2015.

Dimitrova D, **Horak FB**, **Nutt JG**. Postural muscle responses to multidirectional translations in patients with Parkinson's disease. *J Neurophysiol* 91: 489–501, 2004.

Eaton RC, **Emberley DS**. How stimulus direction determines the trajectory of the Mauthner-initiated escape response in a teleost fish. *J Exp Biol* 161: 469–487, 1991.

Fahn S, Elton RL. UPDRS Development Committee. The Unified Parkinson's Disease Rating Scale. In: . *Recent Developments in Parkinson's Disease*. Florham Park, NJ: Macmillan Healthcare Information, 1987, p. 153–163.

Franchignoni F, **Martignoni E**, **Ferriero G**, **Pasetti C**. Balance and fear of falling in Parkinson's disease. *Parkinsonism Relat Disord* 11: 427–433, 2005.

Goetz CG, **Tilley BC**, **Shaftman SR**, **Stebbins GT**, **Fahn S**, **Martinez-Martin P**, **Poewe W**, **Sampaio C**, **Stern MB**, **Dodel R**, **Dubois B**, **Holloway R**, **Jankovic J**, **Kulisevsky J**, **Lang AE**, **Lees A**, **Leurgans S**, **LeWitt PA**, **Nyenhuis D**, **Olanow CW**, **Rascol O**, **Schrag A**, **Teresi JA**, **van Hilten JJ**, **LaPelle N**. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* 23: 2129–2170, 2008.

Govender S, **Dennis DL**, **Colebatch JG**. Axially evoked postural reflexes: influence of task. *Exp Brain Res* 233: 215–228, 2015.

Grillon C, **Ameli R**, **Woods SW**, **Merikangas K**, **Davis M**. Fear-potentiated startle in humans: effects of anticipatory anxiety on the acoustic blink reflex. *Psychophysiology* 28: 588–595, 1991.

Hair JF Jr, **Anderson RE**, **Tatham RL**, **Black WC**. *Multivariate Data Analysis (4th Ed.): With Readings*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1995.

Hely MA, **Morris JGL**, **Reid WGJ**, **Trafficante R**. Sydney Multicenter Study of Parkinson's disease: non-L-dopa-responsive problems dominate at 15 years. *Mov Disord Off J Mov Disord Soc* 20: 190–199, 2005.

Hershenson M. Reaction time as a measure of intersensory facilitation. *J Exp Psychol* 63: 289–293, 1962.

Horak FB, **Henry SM**, **Shumway-Cook A**. Postural Perturbations: New Insights for Treatment of Balance Disorders. *Phys Ther* 77: 517–533, 1997.

Horak FB, **Nashner LM**. Central programming of postural movements: adaptation to altered support-surface configurations. *J Neurophysiol* 55: 1369–1381, 1986.

Hunt AL, **Sethi KD**. The pull test: a history. *Mov Disord Off J Mov Disord Soc* 21: 894–899, 2006.

Jacobs JV, **Dimitrova DM**, **Nutt JG**, **Horak FB**. Can stooped posture explain multidirectional postural instability in patients with Parkinson's disease? *Exp Brain Res Exp Hirnforsch Exp Cerebrale* 166: 78–88, 2005.

Jacobs JV, **Horak FB**, **Van Tran K**, **Nutt JG**. An alternative clinical postural stability test for patients with Parkinson's disease. *J Neurol* 253: 1404–1413, 2006.

Kohfeld DL. Effects of the intensity of auditory and visual ready signals on simple reaction time. *J Exp Psychol* 82: 88–95, 1969.

Kumru H, **Valls-Solé J**. Excitability of the pathways mediating the startle reaction before execution of a voluntary movement. *Exp Brain Res* 169: 427–432, 2006.

Maki BE, **McIlroy WE**. The role of limb movements in maintaining upright stance: the "change-in-support" strategy. *Phys Ther* 77: 488–507, 1997.

McAbee GN. Clobazam-clonazepam combination effective for stimulus-induced falling in hyperekplexia. *J Child Neurol* 30: 91–92, 2015.

McVey MA, **Amundsen S**, **Barnds A**, **Lyons KE**, **Pahwa R**, **Mahnken JD**, **Luchies CW**. The effect of moderate Parkinson's disease on compensatory backwards stepping. *Gait Posture* 38: 800–805, 2013.

Morita H, **Hass CJ**, **Moro E**, **Sudhyadhom A**, **Kumar R**, **Okun MS**. Pedunculopontine Nucleus Stimulation: Where are We Now and What Needs to be Done to Move the Field Forward? *Front Neurol* 5: 243, 2014.

Munhoz R., **Li J-Y**, **Kurtinecz M**, **Piboolnurak P**, **Constantino A**, **Fahn S**, **Lang A.** Evaluation of the pull test technique in assessing postural instability in Parkinson's disease. *Neurol January 13 2004* 62: 125–127, 2004.

Nanhoe-Mahabier W, **Allum JHJ**, **Overeem S**, **Borm GF**, **Oude Nijhuis LB**, **Bloem BR**. First trial reactions and habituation rates over successive balance perturbations in Parkinson's disease. *Neuroscience* 217: 123–129, 2012.

Nieuwenhuijzen PHJA, **Horstink MW**, **Bloem BR**, **Duysens J**. Startle responses in Parkinson patients during human gait. *Exp Brain Res* 171: 215–224, 2006.

Nonnekes J, **Goselink R**, **Weerdesteyn V**, **Bloem BR**. The retropulsion test: a good evaluation of postural instability in Parkinson's disease? *J Park Dis* 5: 43–47, 2015.

Nonnekes J, **Scotti A**, **Oude Nijhuis LB**, **Smulders K**, **Queralt A**, **Geurts ACH**, **Bloem BR**, **Weerdesteyn V**. Are postural responses to backward and forward perturbations processed by different neural circuits? *Neuroscience* 245: 109–120, 2013.

Oliveira AB, **Silva LCCB**, **Coury HJCG**. How do low/high height and weight variation affect upper limb movements during manual material handling of industrial boxes? *Braz J Phys Ther* 15: 494–502, 2011.

Oude Nijhuis LB, **Allum JHJ**, **Borm GF**, **Honegger F**, **Overeem S**, **Bloem BR**. Directional Sensitivity of "First Trial" Reactions in Human Balance Control. *J Neurophysiol* 101: 2802–2814, 2009.

Oude Nijhuis LB, **Allum JHJ**, **Valls-Solé J**, **Overeem S**, **Bloem BR**. First trial postural reactions to unexpected balance disturbances: a comparison with the acoustic startle reaction. *J Neurophysiol* 104: 2704–2712, 2010.

Pai Y-C, **Rogers MW**, **Patton J**, **Cain TD**, **Hanke TA**. Static versus dynamic predictions of protective stepping following waist–pull perturbations in young and older adults. *J Biomech* 31: 1111–1118, 1998.

Perera T, **Tan JL**, **Cole MH**, **Yohanandan SAC**, **Silberstein P**, **Cook R**, **Peppard R**, **Aziz T**, **Coyne T**, **Brown P**, **Thevathasan W**. Balance control systems in Parkinson's disease and the impact of pedunculopontine area stimulation. *Brain* In-press, 2018.

Perera T, **Yohanandan SAC**, **Thevathasan W**, **Jones M**, **Peppard R**, **Evans AH**, **Tan JL**, **McKay CM**, **McDermott HJ**. Clinical validation of a precision electromagnetic tremor measurement system in participants receiving deep brain stimulation for essential tremor. *Physiol Meas* 37: 1516–1527, 2016.

Peterson DS, **Dijkstra BW**, **Horak FB**. Postural motor learning in people with Parkinson's disease. *J Neurol* 263: 1518–1529, 2016.

Ravichandran VJ, **Honeycutt CF**, **Shemmell J**, **Perreault EJ**. Instruction-dependent modulation of the long-latency stretch reflex is associated with indicators of startle. *Exp Brain Res Exp Hirnforsch Exp Cerebrale* 230: 59–69, 2013.

Shemmell J. Interactions between stretch and startle reflexes produce task-appropriate rapid postural reactions. *Front Integr Neurosci* 9, 2015.

Skurvydas A, **Gutnik B**, **Zuoza AK**, **Nash D**, **Zuoziene IJ**, **Mickeviciene D**. Relationship between simple reaction time and body mass index. *HOMO - J Comp Hum Biol* 60: 77–85, 2009.

Smith BA, **Carlson-Kuhta P**, **Horak FB**. Consistency in Administration and Response for the Backward Push and Release Test: A Clinical Assessment of Postural Responses: Consistency of Push and Release Test. *Physiother Res Int* 21: 36–46, 2016.

Sutter K, **Nonnekes J**, **Dibilio V**, **Geurts AC**, **Weerdesteyn V**. Does the StartReact Effect Apply to First-Trial Reactive Movements? *PLOS ONE* 11: e0153129, 2016.

Thevathasan W, **Coyne TJ**, **Hyam JA**, **Kerr G**, **Jenkinson N**, **Aziz TZ**, **Silburn PA**. Pedunculopontine Nucleus Stimulation Improves Gait Freezing in Parkinson Disease: *Neurosurgery* 69: 1248–1254, 2011.

Tresch UA, **Perreault EJ**, **Honeycutt CF**. Startle evoked movement is delayed in older adults: implications for brainstem processing in the elderly. *Physiol Rep* 2, 2014.

Valls-Solé J. Contribution of subcortical motor pathways to the execution of ballistic movements. *Suppl Clin Neurophysiol* 57: 554–562, 2004.

Valls-Sole J, **Rothwell JC**, **Goulart F**, **Cossu G**, **Munoz E**. Patterned ballistic movements triggered by a startle in healthy humans. *J Physiol* 516: 931–938, 1999.

Valls-Solé J, **Solé A**, **Valldeoriola F**, **Muñoz E**, **Gonzalez LE**, **Tolosa ES**. Reaction time and acoustic startle in normal human subjects. *Neurosci Lett* 195: 97–100, 1995.

Visser M, **Marinus J**, **Bloem BR**, **Kisjes H**, **van den Berg BM**, **van Hilten JJ**. Clinical tests for the evaluation of postural instability in patients with parkinson's disease. *Arch Phys Med Rehabil* 84: 1669–1674, 2003.

Woodworth R. *Experimental psychology*. Oxford, England: Holt, 1938.

FIGURES

Figure 1. Set up of instrumented pull test. The instrumented pull test allows an assessor to apply a shoulder-level backward perturbation using a rope and harness **(a)**. The force of the perturbation is recorded using a force gauge **(b)**; the truncal response via a sensor placed at the sternal notch **(c)**; and stepping via sensors on the left and right ankle malleolus **(d)**. The motion tracking system encompasses a processing unit **(e)** which calculates three-dimensional positions of up to four sensors with respect to an electromagnetic transmitter **(f)**. Real time monitoring and feedback is displayed. Auditory stimuli are delivered via headphones. Computerized foot-pedals **(g)** allow quick access to software functions.

Figure 2. Data collected from a representative trial from the instrumented pull test. Vertical dashed lines indicate markers on the time axis. Onset of pull occurs at $\mathbf{t_0}$ with subsequent onset of trunk displacement at $\mathbf{t_1}$. Positive truncal displacement indicates backward movement. The auditory stimulus begins at the falling edge of the sound trigger, within 21 ± 6 ms of peak pull force. Onset of trunk deceleration at $\mathbf{t_2}$ occurs at reversal of peak trunk velocity. The postural response, i.e. truncal reaction time, is defined as the difference between t_2 and t_1.

Figure 3. Raw data (postural response task) from a participant demonstrating single trials associated with the first pull (first trial), normal stimulus at 90 dB (normal) and loud auditory stimulus at 116 dB (LAS). StartReact is demonstrated by quicker reaction times in trunk velocity to the first trial and LAS compared with the normal auditory stimulus. Response magnitude to the postural task is derived from trunk acceleration. The largest response magnitude is demonstrated in the LAS trial, as indicated by the peak of the acceleration curve.

.

TABLES

Trial Type Comparison	Trunk Reaction Time			Trunk Response Magnitude		
	Mean Δ (ms)	95% CI	p-value	Mean Δ ($mm.s^{-2}$)	95% CI	p-value
First vs. Normal	-6.0	-31.1, 19.0	0.692	162	-412, 737	0.497
First vs. Loud	4.2	-21.2, 29.6	0.692	-425	-1008, 158	0.120
Normal vs. Loud	10.2	3.0, 17.5	0.002	-588	-750, -425	< 0.001

Table 1. Mean differences (Δ) between the first pull test trial and subsequent trials with 90 dB (normal) or 116 dB (loud) auditory stimuli for trunk reaction time and response magnitude.

Trial Type Comparison	Step Reaction Time			Step Response Magnitude		
	Mean Δ (ms)	95% CI	p-value	Mean Δ ($mm.s^{-2}$)	95% CI	p-value
First vs. Normal	36.9	4.7, 69.2	0.009	60	17, 103	0.002
First vs. Loud	46.1	13.1, 79.2	0.002	53	9, 97	0.005
Normal vs. Loud	9.2	-3.1, 21.5	0.072	-7	-23, 9	0.315

Table 2. Mean differences (Δ) between the first pull test trial and subsequent trials with 90 dB (normal) or 116 dB (loud) auditory stimuli for step reaction time and response magnitude.

Predictor	Trunk Reaction Time			Trunk Response Magnitude		
	Estimate	95% CI	p-value	Estimate	95% CI	p-value
Peak Force	0.36	0.22, 0.51	< 0.001	0.98	-2.95, 4.91	0.623
Force Rate	-0.01	-0.03, 0.00	0.062	-0.12	-0.47, 0.22	0.486
Height	45.97	-31.16, 123.11	0.233	-708.94	-3362.70, 1944.82	0.587
Weight	-0.17	-0.75, 0.42	0.566	2.08	-18.04, 22.19	0.834

Table 3. Coefficient estimates, 95% confidence intervals (CI), and statistical significance of instrumented pull test predictors resulting from linear mixed models for truncal response.

Predictor	Step Reaction Time			Step Response Magnitude		
	Estimate	95% CI	p-value	Estimate	95% CI	p-value
Peak Force	-0.12	-0.44, 0.19	0.436	1.02	0.55, 1.49	< 0.001
Force Rate	-0.01	-0.04, 0.02	0.575	0.01	-0.03, 0.06	0.528
Height	-64.65	-283.98, 154.69	0.542	240.26	-797.51, 1278.03	0.629
Weight	2.37	0.72, 4.03	0.008	-2.51	-10.56, 5.55	0.518

Table 4. Coefficient estimates, 95% confidence intervals (CI), and statistical significance of instrumented pull test predictors resulting from linear mixed models for step response.