Catch the ruler: concurrent validity and test–retest reliability of the ReacStick measures of reaction time and inhibitory executive function in older people

van Schooten, Kimberley S.; Duran, Lionne; Visschedijk, Maike; Pijnappels, Mirjam; Lord, Stephen R.; Richardson, James; Delbaere, Kim

published in
Aging Clinical and Experimental Research
2019

DOI (link to publisher)
10.1007/s40520-018-1050-6

document version
Publisher's PDF, also known as Version of record

citation for published version (APA)
van Schooten, K. S., Duran, L., Visschedijk, M., Pijnappels, M., Lord, S. R., Richardson, J., & Delbaere, K. (2019). Catch the ruler: concurrent validity and test–retest reliability of the ReacStick measures of reaction time and inhibitory executive function in older people. *Aging Clinical and Experimental Research*, *31*(8), 1147-1154. https://doi.org/10.1007/s40520-018-1050-6

Aging Clinical and Experimental Research (2019) 31:1147–1154
https://doi.org/10.1007/s40520-018-1050-6

ORIGINAL ARTICLE

Catch the ruler: concurrent validity and test–retest reliability of the *ReacStick* measures of reaction time and inhibitory executive function in older people

Kimberley S. van Schooten[1] · **Lionne Duran**[1,2] · **Maike Visschedijk**[1,2] · **Mirjam Pijnappels**[2] · **Stephen R. Lord**[1] · **James Richardson**[3] · **Kim Delbaere**[1]

Received: 15 July 2018 / Accepted: 26 September 2018 / Published online: 9 October 2018

Abstract

Background Reduced cognitive function, particularly executive function (EF), is associated with an increased risk of falling in older people. We evaluated the utility of the *ReacStick* test, a clinical test of reaction time, and inhibitory EF developed, for young athletes, for fall-risk assessment in older people.

Aims To evaluate the psychometric properties of *ReacStick* measures of reaction time and executive functioning in healthy community-dwelling older people.

Methods 140 participants (aged 77 ± 5 years) underwent testing. Two test conditions—simple and inhibitory go/no-go—provided measures of reaction time, recognition load (difference in reaction time between conditions), and go/no-go accuracy. Concurrent validity was evaluated against the conventional tests of reaction time and EF (simple hand reaction time, trail-making test, and Stroop colour test). Discriminant ability was determined for fall-risk factors (age, gender, physiological profile assessment, and fall history). Test–retest reliability after 1 week was evaluated in 30 participants.

Results *ReacStick* reaction time correlated with tests of reaction time and EF, recognition load correlated with inhibitory EF, and go accuracy correlated with reaction time and inhibitory EF. No-go accuracy was not significantly correlated with any of the reaction time and EF tests. Test–retest reliability was good-to-excellent (ICC > 0.6) for all the outcomes. *ReacStick* reaction time discriminated between groups based on age, recognition load between genders, and no-go accuracy between retrospective fallers and non-fallers.

Discussion An unavoidable time pressure may result in complementary information to the traditional measures.

Conclusions The *ReacStick* is a reliable test of reaction time and inhibitory EF in older people and could have value for fall-risk assessment.

Keywords Task switching · Cognition · Go–no-go · Accidental falls · Aged

Background

Research has shown that gait is not just a rhythmic and automatic process but requires cognitive resources [1–3]. These demands on cognition increase with age and the complexity of the task being performed. As such, adequate cognitive function is essential for maintaining stable gait and reducing fall risk in older people [4, 5]. More specifically, executive limitations have been associated with falls and fall-related injuries in older people [6, 7]. Executive function refers to skills for achieving goal-directed behaviours [8], and include the ability to inhibit pre-potent but inappropriate responses, selectively attend to relevant environmental factors, and plan and strategize actions to react properly to task-relevant

✉ Kim Delbaere
k.delbaere@neura.edu.au

[1] Neuroscience Research Australia, University of New South Wales, Sydney, Australia

[2] Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Vrije Universiteit, Amsterdam Movement Sciences, Amsterdam, The Netherlands

[3] Department of Physical Medicine and Rehabilitation, University of Michigan, Ann Arbor, USA

stimuli [9, 10]. Executive functions include basic cognitive processes [11, 12] which can be broken down into three core executive functions: working memory, set-shifting (also coined cognitive or mental flexibility), and inhibition; the latter including self-control (behavioural inhibition) and interference control (attentional control and cognitive inhibition) [12]. Higher order executive functions, such as decision-making and planning, rely on these core executive functions [13, 14]. In the context of falls, adequate working memory may assist a person to monitor their ongoing movements and adjust these to the environment, set-shifting may assist a person to focus on walking in a busy environment (and switch between tasks as required), and inhibition allows a person to block out task-irrelevant information and 'change' their step mid-air to avoid a slip. Together, adequate and simultaneous use of these core executive functions allows a person to select the correct motor response after being perturbed.

The effect of attentional and executive limitations on balance and fall risk has been extensively studied using dual-task experiments where concurrent cognitive tasks are performed [8]. Alternatively, a more isolated (or single task) approach has been used to associate balance impairments and fall risk with the tests of executive function, such as the trail-making test, verbal fluency test (although debatable if this test assesses executive function [15]), clock-drawing test, forward and backward digit span tests, and Stroop tests. Several studies suggest that tests that encompass inhibitory as well as other executive functions are associated with falls in older people [16, 17]. The complexity of executive function and the different domains it covers make it difficult to distinguish cognitive processes that actually limit performance on these tests. For example, the trail-making and clock-drawing tests rely on fine-motor skills, the verbal fluency test on language, and the Stroop test on colour identification and reading skills. To fully understand the impact of high-order cognitive processing on daily life functioning and fall risk, it is important to use a task-specific executive function measure that rely minimally on other functions. To date, a lack of specific and time-contingent tests renders it uncertain which executive functions contribute to balance and gait impairments and increased fall risk in older people.

Eckner and coworkers [18] developed the *ReacStick* test as a clinical measure of simple reaction time and short-latency inhibitory processing efficiency using a go/no-go paradigm. The test requires a person to catch (or not catch) a falling stick after it is released by an examiner. It is a significant advance to more traditional executive function tests, because it adds an unavoidable time pressure that demands a quick response within approximately 390 ms [18], which markedly limits the possibility of a strategy that trades speed for accuracy. Moreover, this test does not rely on other factors such as language or fine-motor skills. The previous studies utilizing the *ReacStick* test showed that the simple reaction time component (SRT) is highly motivating and clinically feasible [19]. In addition, *ReacStick* SRT discriminated between athletes with and without sport-related concussion and correlated strongly with the ability to protect the head and face from a projectile threat [20, 21]. Pilot work with the *ReacStick* go/no-go, or recognition reaction time (RRT), component showed that latencies were significantly prolonged compared to SRT latency, with the majority of this delay related to pre-movement time rather than digit closure, indicating that the delay was likely of central origin [18].

The aim of the present study was to establish the psychometric properties of the *ReacStick* test by means of construct validity for reaction time and (inhibitory) executive functions, test–retest reliability, and discriminant ability against known fall-risk factors in a sample of community-dwelling older people. We hypothesized that performance on the *ReacStick* simple reaction time task would be associated with cognitive processing speed and that performance on the go/no-go task would be associated with tests of inhibitory executive functions. We further anticipated that both tests would be associated with fall risk in older people.

Methods

Participants

One hundred and forty older people participated in this study. Participants were community-dwelling, aged 70 years or older, lived in the Sydney metropolitan area and took part in the *StandingTall* randomized-controlled trial [22] conducted at Neuroscience Research Australia. Participants were English-speaking and able to walk without the use of a walking aid in their home. Exclusion criteria were: having (1) an unstable or acute medical condition precluding exercise, (2) a progressive neurological disease, and (3) a cognitive impairment as indicated by a Pfeiffer Short Portable Mental Status Questionnaire score < 8. The study protocol was approved by the UNSW Ethics Committee (reference HC#14/266).

Participants were on average 77.2 (SD 5.4) years old and 92 out of 140 (65.7%) were female. They had an average of 14.2 (3.9) years of education. One hundred and thirteen participants (80.7%) reported to be in very good health and 64 participants (45.7%) reported a fall in the prior 12 months.

ReacStick assessment of reaction time and inhibition

The *ReacStick* is a rigid, lightweight bar with a built-in accelerometer, timing circuit and LED light (Fig. 1). This

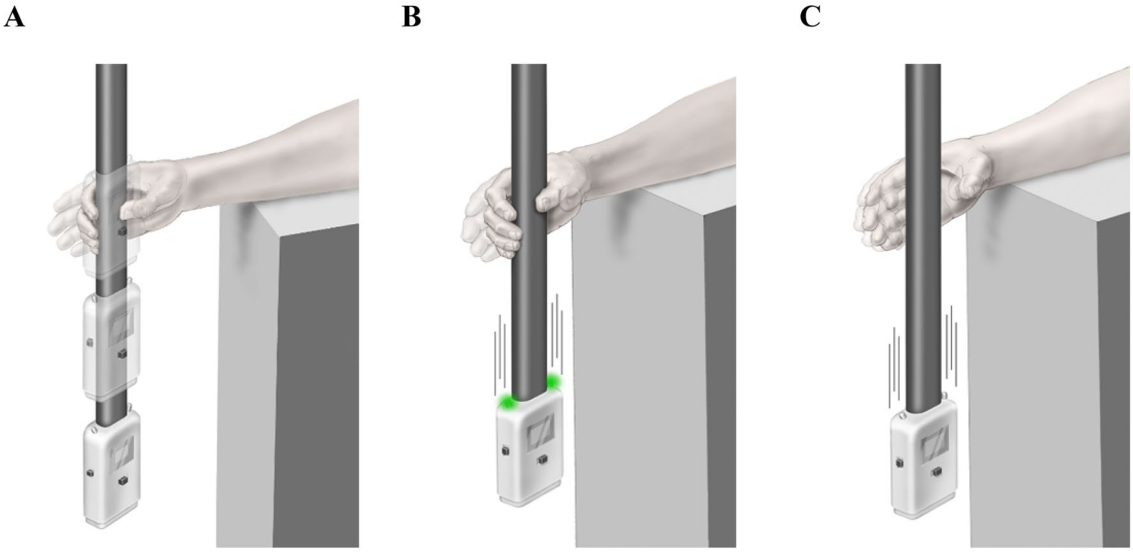

Fig. 1 *ReacStick* assessment of simple reaction time (SRT, **a**) and recognition reaction time (RRT), where the device is caught when the participant is instructed to catch the device when the green light turns on (RRT ON, **b**), and to inhibit the urge to catch the device when the green light does not turn on (RRT OFF, **c**)

device was used to measure simple reaction time (SRT) and recognition (go/no-go) reaction time (RRT) [21]. During testing, participants were seated with the forearm of their dominant arm on a desk at 90° elbow flexion and the ulnar styloid in contact with the edge of the desk (Fig. 1). To assess SRT, the *ReacStick* device was lifted vertically from its upper end by the examiner, keeping the device's housing in the hand of the participant without contact. After a random delay between 2 and 8 s, the examiner released the *ReacStick*. Participants were instructed to catch the falling device as quickly as possible by closing their hand. Five SRT practice and ten SRT test trials were performed. SRT was obtained from the electronic display of the device, and was defined as the time from initial device acceleration due to its release to initial deceleration due to participant hand contact [18].

The RRT trials were similar to the SRT trials, but included a go/no-go paradigm. During the RRT trials, participants were instructed to catch the *ReacStick* only when green LED lights on top of the housing illuminated at the instant of release during a random 50% of trials; these "go" trials referred to as the RRT ON trials. Participants were further instructed to let the *ReacStick* fall to the ground on the random 50% of trials when the lights did not illuminate; these "no-go" trials referred to as RRT OFF trials. Five RRT practice trials and twenty RRT test trials were performed. RRT was measured as the time required for catching the device for the correctly performed ON trials. ON and OFF accuracy was expressed as the percentage of correct responses to ON and OFF trials, respectively. Recognition load, the additional time used to process and respond in RRT ON trials, was defined as the difference in time between RRT ON and SRT. We repeated the *ReacStick* test in a random subsample of 30 participants 1 week later to investigate its test–retest reliability.

Demographic characteristics and fall-risk measures

Age, gender, and fall history, defined as having had a fall in the past year, were obtained during a structured interview. The Physiological Profile Assessment (PPA; Neuroscience Research Australia, Sydney NSW, Australia) was used to estimate physiological fall risk [23]. The PPA includes five tests of different sensorimotor functions: visual contrast sensitivity with the Melbourne edge test, proprioception measured with the lower limb matching-task, knee extension strength measured in the dominant leg, simple hand reaction time, and postural sway while standing on a foam rubber mat.

Measures of reaction time and inhibition

Concurrent validity of performance on the *ReacStick* test was assessed against a simple hand reaction time test, the trail-making test, and the Victoria Stroop colour test. Simple hand reaction time was measured as part of the PPA, with a light as stimulus and a finger press as response, to assess processing speed.

The trail-making test (TMT) was used as a measure of visual search and executive function, particularly psychomotor speed, set-shifting, and working memory [24, 25]. In TMT part A, the participant is asked to connect numbered circles on an electronic tablet by tapping the targets in sequential order as fast as possible. The instruction on TMT part B is similar to TMT part A except that the circles contain numbers and letters, and the participant is asked to connect the circles in alternating numeric and alphabetic sequence (i.e., 1, A, 2, B, etc). Duration of TMT part B minus TMT part A (TMT B-A) provides further insight into set-shifting.

The Victoria Stroop colour test was used as a conflict resolution task that includes inhibition [24]. In condition 1 (or part D), participants are asked to state the colour of a dot presented on a computer screen, and in condition 2 (or part C), participants are asked to state the colour of the ink rather than verbalise the written words. In the current study, we used the number of errors during the colour-word interference task (intrusions) and the ratio of colour-word interference and colour only tasks (inefficiency).

Statistical analysis

SPSS for Windows (version 24, SPPS, Inc., Chicago, IL, USA) was used for the statistical analyses. Normality of the data was checked using the Shapiro–Wilk test and visual inspection, which revealed that ON and OFF accuracy (both percentage-scores) did not follow a normal distribution. We, therefore, report the median and IQR for these measures and the mean and SD for all others. We first determined correlations within *ReacStick* measures to identify whether all they assessed similar constructs. Second, we assessed concurrent validity of the *ReacStick* measures with the hand reaction time, trail-making, and Stroop tests. We expected that SRT would associate with central processing speed as assessed by hand reaction time and trail-making test. We further expected that recognition load (difference between RRT and SRT) would associate with inhibitory executive function as assessed by the Stroop test, and less so with working memory and set-shifting as assessed by the trail-making test. Since ON and OFF accuracy did not follow a normal distribution and since Spearman and Pearson correlations provided similar results for the normally distributed measures, we report Spearman correlations for all measures. A correlation between 0.1 and 0.29 was considered weak, between 0.3 and 0.49 moderate, and between 0.5 and 1.0 strong [26]. Third, we assessed test–retest reliability of the *ReacStick* measures using the one-way random effects model. Intra-class Correlation Coefficient (ICC 2.1) between the scores obtained during the baseline assessment and the scores obtained 1 week later. We also report ICC for ON and OFF accuracy, because the obtained ICCs were comparable to Spearman correlations and Wilcoxon signed-rank tests did not indicate a significant difference in measures between both weeks. An ICC below 0.4 was considered poor, between 0.4 and 0.59 fair, between 0.6 and 0.74 good, and between 0.75 and 1.00 excellent [27]. Fourth, we determined discriminant ability for fall-risk factors using independent *T* tests or Mann–Whitney *U* test. Measures of the *ReacStick* were compared to groups categorised by gender, age (median cut point), PPA score, and history of falls, since these factors have been associated with falls and fall risk.

Results

Participants had a mean score of 0.9 (SD 0.8) on the PPA, indicating a low-to-moderate physiological fall risk. They had an average hand reaction time of 238.3 (SD 41.3) ms and took 31.7 (SD 10.5) s to complete TMT part A and 99.8 (SD 47.0) s to complete TMT part B, resulting in a TMT B-A of 68.2 (SD 47.0) s. Participants had a mean Stroop inefficiency score of 2.0 (SD 0.9) and had an average of 4.0 intrusions (Inter-quartile Range, IQR 5.0) on the Stroop test.

Inter-item correlations

All participants were able to complete the *ReacStick* test. The mean scores for the SRT and RRT ON trials were 172.6 (SD 16.8) ms and 247.3 (SD 32.6) ms, respectively ($p < 0.01$), resulting in an average recognition load of 74.7 (SD 31.7) ms (Table 1). Participants had a median ON accuracy of 80% (IQR 20%) and OFF accuracy of 40% (IQR 40%). Those who took longer to catch the *ReacStick* during the RRT ON trials had a lower ON accuracy and higher OFF accuracy ($\rho = -0.768$, $p < 0.001$ and $\rho = 0.359$, $p < 0.001$, respectively), suggesting a delayed response under both conditions. The inter-item correlations between the *ReacStick* measures ranged from weak to strong (see Table 1).

Concurrent validity

Correlations between *ReacStick* and other measures of simple reaction time and executive functions were significant but weak, ranging between -0.24 and 0.22 (Table 2). A longer SRT mean correlated with a longer hand reaction time, longer time taken to complete the trail-making and higher Stroop inefficiency, and more intrusions on the Stroop test. The Bland–Altman plot describes the level of agreement between simple reaction times and SRT and shows that reaction times were faster during the SRT when compared to hand reaction time (Fig. 2). A higher RRT ON mean correlated with longer hand reaction time. The recognition load correlated with the Stroop inefficiency and Stroop intrusions, where higher recognition load was associated with a better Stroop performance. RRT ON accuracy correlated with quicker hand reaction time and a higher Stroop inefficiency. RRT OFF accuracy did not correlate significantly with any of the concurrent validity measures.

Table 1 Group averages and internal consistency of the ReacStick measures ($N = 140$)

	Mean (SD) or median [IQR]	Spearman correlation coefficients			
		RRT ON mean	Recognition load	RRT ON accuracy	RRT OFF accuracy
SRT mean (ms)	172.6 (16.8)	0.31**	− 0.22*	− 0.12	0.04
RRT ON mean (ms)	247.3 (32.6)		0.83**	− 0.77**	0.36**
Recognition load (ms)	74.7 (31.7)			− 0.71**	0.33**
RRT ON accuracy (%)	80 [20]				− 0.13
RRT OFF accuracy (%)	40 [40]				

SD, standard deviation; IQR, inter-quartile range; SRT, simple reaction time, RRT, recognition reaction time

*$p < 0.05$; **$p < 0.01$

Table 2 Spearman correlations between ReacStick measures and concurrent validity measures ($N = 140$)

	SRT mean	RRT ON mean	Recognition load	RRT ON accuracy	RRT OFF accuracy
Hand reaction time	0.168*	0.192*	0.114	− 0.180*	0.152
TMT A	0.215*	− 0.030	− 0.156	− 0.017	− 0.069
TMT B	0.208*	0.080	− 0.037	− 0.116	− 0.098
TMT B-A	0.182*	0.096	0.003	− 0.120	− 0.104
Stroop inefficiency	0.200*	− 0.119	− 0.240**	0.166*	− 0.048
Stroop intrusions	0.198*	− 0.117	− 0.229**	0.110	− 0.117

SRT, simple reaction time, RRT, recognition reaction time

*$p < 0.05$; **$p < 0.01$

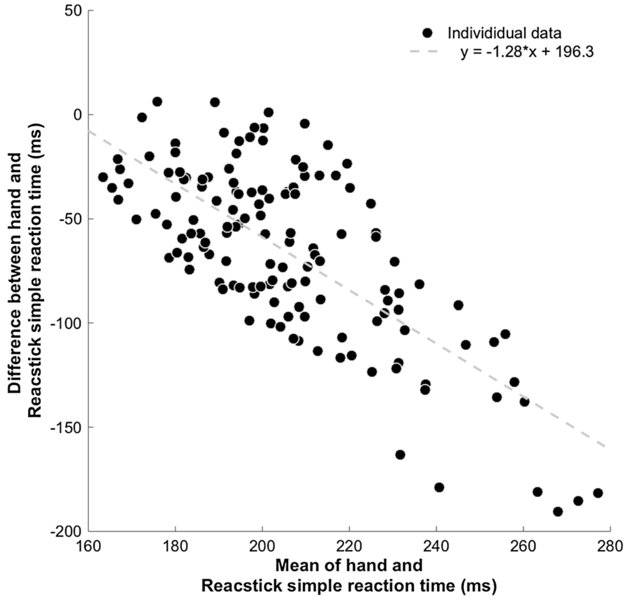

Fig. 2 Bland–Altman plot of simple hand reaction time and *ReacStick* simple reaction time. People with a slower reaction time averaged over the two tests, generally, have a quicker *ReacStick* reaction time compared to their hand reaction time

Test–retest reliability

Table 3 Test–retest reliability ($N = 30$)

	ICC	95% CI
SRT mean	0.98	0.95–0.99
RRT ON mean	0.72	0.49–0.85
Recognition load	0.60	0.32–0.79
RRT ON accuracy	0.63	0.36–0.81
RRT OFF accuracy	0.72	0.49–0.86

ICC, intra-class coefficient; CI, confidence interval; SRT, simple reaction time, RRT, recognition reaction time

Test–retest reliability was excellent for SRT mean (ICC = 0.98), and good for RRT mean, recognition load, RRT ON accuracy, and RRT OFF accuracy (ICC ≥ 0.6) (Table 3).

Discriminant ability

Table 4 presents differences in *ReacStick* measures for groups based on demographic characteristics and fall-risk indicators. SRT discriminated between age groups, with participants older than 76 years showing significantly longer SRT mean durations ($p = 0.047$). RRT and recognition load discriminated between men and women, with significantly longer RRT mean durations and a higher recognition

Table 4 ReacStick mean (SD) measures for subgroups based on fall-risk factors ($N = 140$)

Variables	Cut point	*N*	SRT mean	RRT ON mean	Recognition load	RRT ON accuracy	RRT OFF accuracy
Age	≤76	72	169.8 (16.1)*	247.0 (30.6)	77.2 (29.2)	79.6 (14.8)	41.8 (21.4)
	>76	68	175.5 (17.2)*	247.6 (34.7)	72.1 (34.1)	78.7 (16.8)	38.2 (23.4)
Gender	Male	48	172.0 (16.7)	237.8 (36.7)*	65.8 (35.3)*	81.9 (15.9)	35.6 (23.8)
	Female	92	173.0 (17.0)	252.2 (29.2)*	79.4 (28.7)*	77.7 (15.6)	42.4 (21.4)
PPA score	Low risk	55	170.1 (17.0)	247.7 (28.1)	77.6 (27.6)	81.1 (13.7)	43.5 (19.5)
	High risk	85	174.2 (16.6)	247.1 (35.3)	72.9 (34.1)	77.9 (16.9)	37.9 (24.0)
Fall history	0	94	172.8 (15.9)	248.7 (32.9)	75.9 (30.8)	77.9 (16.3)	43.9 (23.6)**
	≥1	46	172.2 (18.8)	244.4 (32.1)	72.2 (33.7)	81.7 (14.3)	32.2 (17.4)**

SRT, simple reaction time, RRT, recognition reaction time

*$p < 0.05$; **$p < 0.01$

load in women compared to men ($p = 0.012$ and $p = 0.016$, respectively). None of the *ReacStick* measures discriminated between people with a high or low risk for falls based on their PPA score. RRT OFF accuracy discriminated between people with and without a history of falls. Participants who experienced one or more falls in the past 12 months were significantly less accurate in the RRT OFF trials ($p = 0.001$) compared to people who did not experience one or more falls in the past 12 months.

Discussion

The current study evaluated the psychometric properties of *ReacStick* measures of reaction time and executive functioning in community-dwelling older people. Test–retest reliability over a week time interval was good-to-excellent. Simple and go/no-go *ReacStick* measures showed weak-to-strong correlations with each other. Participants who took longer on the SRT also took longer on the RRT ON trials. Participants who were slower on the RRT ON trails showed a higher recognition load and lower RRT ON accuracy, but higher RRT OFF accuracy. Nevertheless, all ten participants with RRT OFF accuracy of 100% had RRT ON accuracies exceeding 40%, suggesting that they did understand the task at hand. Participants who were accurate on both RRT ON and OFF trials demonstrated good selective attention, had the capacity to withhold the reactive urge to catch the falling device, and process the go/no-go stimuli and the ability to instigate the correct response. These findings suggest that the *ReacStick* measures reflect overlapping but different aspects; which could be further discerned from our construct validity results.

To test the concurrent validity of the *ReacStick* measures, we assessed their associations with central processing speed as assessed by hand reaction time and trail-making test, inhibitory executive function as assessed by the Stroop test, and working memory and set-shifting as assessed by the trail-making test. We found that simple hand reaction time and mean *ReacStick* reaction time, both measures of processing speed, were only weakly correlated. We observed quicker reactions on the SRT compared to the simple hand reaction time (172.6 SD 16.8 vs. 238.3 SD 41.3 ms, respectively; $p < 0.001$), consistent with prior work which found SRT to be faster than a simple keyboard spacebar reaction time test [28, 29]. This was also confirmed on an individual level, where we observed that people who had slower reaction time were relatively quicker on the SRT (Fig. 2). Potential explanations for this finding could be that the falling *ReacStick* invokes an involuntary or automated reaction to arrest falling objects, and the possibility that central visual processing occurs more quickly for motion compared to identification of an object [29, 30]. As such, SRT may provide the clinician with better insight into "real-world" response latencies than the conventional screens.

We further found that a higher recognition load was associated with a better performance on the Victoria Stroop colour test. This finding suggests that people who are able to inhibit their initial response perform better on both tasks. However, a higher recognition load seems to affect their RRT ON accuracy negatively ($\rho = -0.17$ with recognition load, and $\rho = 0.17$ with Stroop inefficiency), which demonstrates that there is an optimal response time between successful inhibition and being too late to catch the device. RRT ON and OFF accuracies did not correlate with Stroop intrusions as hypothesized, suggesting that motor and verbal response inhibition are distinct processes [31]. Overall, the weak correlations between the *ReacStick* and tests of executive functions suggest that it may be complementary to the traditional measures; however, future studies are required for the validation of this notion.

Finally, we assessed whether the *ReacStick* measures were sensitive to demographic characteristics and fall-risk factors. We found that mean SRT discriminated between age groups, confirming that reaction time is generally higher in older old people [32, 33]. We further found that mean RRT and recognition load discriminated between genders with men being slightly faster than women, a finding that has been reported before [33]. None of the *ReacStick* measures discriminated between people with high and low physiological fall risk according to the PPA, which could indicate that they provide complementary information. RRT OFF accuracy discriminated between people with and without the previous falls and seems to be a promising measure for fall risk. This is strengthened by findings of a recent pilot study where RRT ON + OFF accuracy was found to be more than two standard deviations higher in two individuals with diabetic peripheral neuropathy who experienced injurious falls, compared to the 17 individuals who did not [34]. Larger studies are required to confirm if the *ReacStick*, indeed, provides insight into fall and fall-related injury risk.

ReacStick measurements have clear potential for clinical application. One aspect that makes the *ReacStick* test particularly useful in the clinic is its "obviousness". When patients are demonstrably slow catching the *ReacStick* during SRT testing in the exam room, it is apparent not just to the clinician but to the patient and their family members. Similarly, when patients cannot make the correct decision during RRT evaluation, it is obvious to all that the patient has a problem with quick decision-making and action. These shared observations then allow greater weight to clinician recommendations to limit driving, avoid hazardous walking situations, or use an assistive device. When applied to daily life situations, problems with rapid decision-making could indicate how well a person is able to avoid a potential hazard, cope with unexpected perturbations, and change an action in a split second; each of which could prevent a fall. Recently, Verghese and colleagues used fNIRS to reveal that high-functioning older people at risk of falling have increased brain activity in the prefrontal cortex during cognitively demanding assessments [35]. The authors further suggest that the traditional executive function measures are not able to capture these deficits unless a brain scan assessment, such as fNIRS, is used in conjunction. We hypothesize that the *ReacStick* could capture this rapid cortical inhibition better than the traditional tests because of the unavoidable time pressure, but this obviously needs further exploration.

We acknowledge that this study has certain limitations. First, there is currently no validated response inhibition measure that could be used as a gold standard measure to compare to the *ReacStick* measures. Similarly, while TMT is an accepted measure of working memory [24, 25], there is still debate on what constitutes a good task that combines 'processing' (working) and 'storage' (memory). Second, we did not correct for multiple testing which may have led to type 1 errors. However, considering the large sample and exploratory nature of the current study, we feel that it was warranted. Third, we only evaluated falls of the participants over the past year. This was evaluated through a questionnaire and relies on the memory of the participants and is, therefore, subjected to recall bias. Future research should use prospective measures to investigate whether the *ReacStick* measures could predict falls in older people. Finally, participants were overall in a very good health, with a relatively small age range (70–91 years). As such, the study findings may not be generalizable to frailer older people.

In conclusion, the *ReacStick* is a reliable test of processing speed and inhibitory EF in older people. Based on the high correlations among measures and good test–retest reliability, we identify SRT mean, recognition load, and OFF (no-go) accuracy as promising measures for prospective studies. These measures discriminated between groups based on age, gender, and fall history. Future studies are needed to determine the predictive value of the *ReacStick* towards future falls.

Acknowledgements We like to thank our participants for participating, and research assistants Jessica Chow, Garth McInerney, Ashton May, Lillian Miles, Linda Pickett, Daniel Steffens, and Ashley Woodbury for their help with participant recruitment and assessments.

Funding This work was supported by the Human Frontier Science Program (HFSP long-term fellowship number LT001080/2017); the Dutch Organisation for Scientific Research (NWO VIDI Grant number 91714344); the Newman Family Foundation; the Australian National Health and Medical Research Council (NHMRC project Grant number 1084739).

Compliance with ethical standards

Conflict of interest The authors report no conflicts of interest.

Statement of human and animal rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Springer S, Giladi N, Peretz C et al (2006) Dual-tasking effects on gait variability: the role of aging, falls, and executive function. Mov Disord 21:950–957. https://doi.org/10.1002/mds.20848
2. Yogev-Seligmann G, Hausdorff JM, Giladi N (2008) The role of executive function and attention in gait. Mov Disord 23:329–342. https://doi.org/10.1002/mds.21720
3. Lord S, Galna B, Verghese J et al (2013) Independent domains of gait in older adults and associated motor and nonmotor attributes: validation of a factor analysis approach. J Gerontol Ser A Biol Sci Med Sci 68:820–827. https://doi.org/10.1093/gerona/gls255
4. Herman T, Mirelman A, Giladi N et al (2010) Executive control deficits as a prodrome to falls in healthy older adults: a prospective study linking thinking, walking, and falling. J Gerontol Ser A Biol

Sci Med Sci 65:1086–1092. https://doi.org/10.1093/gerona/glq077

5. Mirelman A, Herman T, Brozgol M et al (2012) Executive function and falls in older adults: new findings from a five-year prospective study link fall risk to cognition. PLoS One 7:e40297. https://doi.org/10.1371/journal.pone.0040297
6. Delbaere K, Kochan NA, Close JC et al (2012) Mild cognitive impairment as a predictor of falls in community-dwelling older people. Am J Geriatr Psychiatry 20:845–853. https://doi.org/10.1097/JGP.0b013e31824afbc4
7. Hsu CL, Nagamatsu LS, Davis JC et al (2012) Examining the relationship between specific cognitive processes and falls risk in older adults: a systematic review. Osteoporos Int 23:2409–2424. https://doi.org/10.1007/s00198-012-1992-z
8. Montero-Odasso M, Verghese J, Beauchet O et al (2012) Gait and cognition: a complementary approach to understanding brain function and the risk of falling. J Am Geriatr Soc 60:2127–2136. https://doi.org/10.1111/j.1532-5415.2012.04209.x
9. Verbruggen F, Logan GD (2009) Models of response inhibition in the stop-signal and stop-change paradigms. Neurosci Biobehav Rev 33:647–661. https://doi.org/10.1016/j.neubiorev.2008.08.014
10. Wecker NS, Kramer JH, Wisniewski A et al (2000) Age effects on executive ability. Neuropsychology 14:409–414
11. Diamond A (2013) Executive functions. Annu rev Psychol 64:135–168
12. Miyake A, Friedman NP, Emerson MJ et al (2000) The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: a latent variable analysis. Cogn Psychol 41:49–100. https://doi.org/10.1006/cogp.1999.0734
13. Collins A, Koechlin E (2012) Reasoning, learning, and creativity: frontal lobe function and human decision-making. PLoS Biol 10:e1001293. https://doi.org/10.1371/journal.pbio.1001293
14. Lunt L, Bramham J, Morris RG et al (2012) Prefrontal cortex dysfunction and ‘Jumping to Conclusions’: bias or deficit? J Neuropsychol 6:65–78. https://doi.org/10.1111/j.1748-6653.2011.02005.x
15. Whiteside DM, Kealey T, Semla M et al (2016) Verbal fluency: language or executive function measure? Appl Neuropsychol Adult 23:29–34. https://doi.org/10.1080/23279095.2015.1004574
16. Anstey KJ, Wood J, Kerr G et al (2009) Different cognitive profiles for single compared with recurrent fallers without dementia. Neuropsychology 23:500
17. Schoene D, Delbaere K, Lord SR (2017) Impaired response selection during stepping predicts falls in older people—a cohort study. J Am Med Dir Assoc 18:719–725
18. Eckner JT, Richardson JK, Kim H et al (2012) A novel clinical test of recognition reaction time in healthy adults. Psychol Assess 24:249–254. https://doi.org/10.1037/a0025042
19. Eckner JT, Chandran S, Richardson JK (2011) Investigating the role of feedback and motivation in clinical reaction time assessment. PM R J Injury Funct Rehabil 3:1092–1097. https://doi.org/10.1016/j.pmrj.2011.04.022
20. Eckner JT, Kutcher JS, Broglio SP et al (2014) Effect of sport-related concussion on clinically measured simple reaction time. Br J Sports Med 48:112–118. https://doi.org/10.1136/bjsports-2012-091579
21. Eckner JT, Lipps DB, Kim H et al (2011) Can a clinical test of reaction time predict a functional head-protective response? Med Sci Sports Exerc 43:382–387. https://doi.org/10.1249/MSS.0b013e3181f1cc51
22. Delbaere K, Valenzuela T, Woodbury A et al (2015) Evaluating the effectiveness of a home-based exercise programme delivered through a tablet computer for preventing falls in older community-dwelling people over 2 years: study protocol for the Standing Tall randomised controlled trial. BMJ Open 5:e009173. https://doi.org/10.1136/bmjopen-2015-009173
23. Lord SR, Menz HB, Tiedemann A (2003) A physiological profile approach to falls risk assessment and prevention. Phys Ther 83:237–252
24. Lezak MD, Howieson DB, Loring DW et al (2004) Neuropsychological assessment. Oxford University Press, USA
25. Sanchez-Cubillo I, Perianez JA, Adrover-Roig D et al (2009) Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. J Int Neuropsychol Soc 15:438–450. https://doi.org/10.1017/S1355617709090626
26. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd. Erlbaum, Hillsdale
27. Cicchetti D (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instrument in psychology. Psychol Assess 6:284–290
28. Eckner JT, Whitacre RD, Kirsch NL et al (2009) Evaluating a clinical measure of reaction time: an observational study. Percept Motor Skills 108:717–720
29. Montare A (2010) The simplest chronoscope II: reaction time measured by meterstick versus machine. Percept Mot Skills 111:819–828
30. Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15:20–25
31. Bari A, Robbins TW (2013) Inhibition and impulsivity: behavioral and neural basis of response control. Progress Neurobiol 108:44–79
32. Salthouse TA (1996) The processing-speed theory of adult age differences in cognition. Psychol Rev 103:403
33. Der G, Deary IJ (2006) Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. Psychol Aging 21:62
34. Richardson JK, Eckner JT, Allet L et al (2017) Complex and simple clinical reaction times are associated with gait, balance, and major fall injury in older subjects with diabetic peripheral neuropathy. Am J Phys Med Rehabil 96:8
35. Verghese J, Wang C, Ayers E et al (2017) Brain activation in high-functioning older adults and falls: Prospective cohort study. Neurology 88:191–197. https://doi.org/10.1212/WNL.0000000000003421