Medical Engineering and Physics 000 (2016) 1–8

Contents lists available at ScienceDirect

Medical Engineering and Physics

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

Technical note

Accuracy and re-test reliability of mobile eye-tracking in Parkinson's disease and older adults

S. Stuart*, L. Alcock, A. Godfrey, S. Lord, L. Rochester, B. Galna

Institute of Neuroscience/Newcastle University Institute for Ageing, Clinical Ageing Research Unit, Newcastle University, United Kingdom

ARTICLE INFO

Article history:
Received 9 May 2015
Revised 29 October 2015
Accepted 13 December 2015
Available online xxx

Keywords:
Parkinson's disease
Mobile eye-tracking
Accuracy
Reliability
Saccades
Walking

ABSTRACT

Mobile eye-tracking is important for understanding the role of vision during real-world tasks in older adults (OA) and people with Parkinson's disease (PD). However, accuracy and reliability of such devices have not been established in these populations. We used a novel protocol to quantify accuracy and reliability of a mobile eye-tracker in OA and PD.

A mobile eye-tracker (Dikablis) measured the saccade amplitudes of 20 OA and 14 PD on two occasions. Participants made saccades between targets placed 5°, 10° and 15° apart. Impact of visual correction (glasses) on saccadic amplitude measurement was also investigated in 10 OA.

Saccade amplitude accuracy (median bias) was −1.21° but a wide range of bias (−7.73° to 5.81°) was seen in OA and PD, with large vertical saccades (15°) being least accurate. Reliability assessment showed a median difference between sessions of <1° for both groups, with poor to good relative agreement (Spearman *rho*: 0.14 to 0.85). Greater accuracy and reliability was observed in people without visual correction.

Saccade amplitude can be measured with variable accuracy and reliability using a mobile eye-tracker in OA and PD. Human, technological and study-specific protocol factors may introduce error and are discussed along with methodological recommendations.

1. Introduction

Eye-tracking provides data regarding the acquisition of visual information, which is crucial for the safe and effective performance of many real-world activities. Eye-tracking devices have become increasingly popular for investigating visual deficits in people with Parkinson's disease (PD) and older adults (OA) [1,2]. Previous eye-tracking studies have typically measured visual activity in static laboratory settings [3]. More recently, mobile eye-tracking devices have allowed researchers to investigate the influence of both PD and ageing on visual exploration during real-world activities such as walking and obstacle crossing [1,2]. Both mechanistic and clinical research requires accurate and reliable devices. However, a recent review [1] highlighted that previous studies do not report the accuracy or reliability of their eye-tracking devices. This is likely due to a lack of 'gold-standard' device or protocol for comparison. As such, there is sparse information regarding the psychometric properties of mobile eye-tracking devices in people with PD and OA.

Previous studies [4–7] have evaluated reliability of static eye-tracking devices in various populations, measuring saccades for specific phenomena using highly specialised protocols. For example, Farzin et al. [7] reported that their static eye-tracker (Tobii, T120, 300 Hz) was reliable in reporting number and duration of fixations, and pupillary response during a seated picture-viewing protocol in Fragile-X syndrome patients and controls. Similarly, other studies have assessed reliability of eye-movement characteristics measured with static devices but focus on specific assessments such as anti- or pro-saccade tests [4,5,8], and attribute reliability differences to disease-related influences rather than the device [4]. Results of these highly specialised protocols are not easily generalised, highlighting the need for a standardised protocol.

A previous study reported the accuracy of a desk-mounted Tobii eye-tracker (TX300, 300 Hz) was 0.5° [9] when participants walked on a treadmill and look at targets on a screen at various locations. The static device had a high sampling-frequency (300 Hz) and accounted for head movement as long as participants stayed within 200 cm of the screen. As such, the results may not apply to head-mounted mobile eye-tracking devices which capture at lower frequencies (i.e. 50–60 Hz) but do not require movement to be restricted [10].

---

* Corresponding author. Tel.: +44191 208 1242.
E-mail address: sam.stuart@newcastle.ac.uk (S. Stuart).

http://dx.doi.org/10.1016/j.medengphy.2015.12.001
1350-4533/

Our previous work [11] has shown that mobile eye-trackers can accurately detect saccades, however little is known about the accuracy or reliability of specific saccade characteristics (e.g. amplitude) recorded via mobile eye-trackers during static or dynamic tasks [1]. This is important as such characteristics can inform disease-related impairment. This study aimed to evaluate accuracy and reliability of a mobile eye-tracker in measurement of saccade amplitude in people with PD and OA when sitting, standing and walking. We developed a simple protocol using visual targets placed at set distances, which could be used to evaluate other devices across different populations.

2. Materials and methods

2.1. Participants

Fourteen people with PD were recruited through local Movement Disorders clinics along with 20 age-matched OA through local advertisements.

Inclusion criteria for all participants were: ≥50 years, normal or corrected-to-normal vision (<18/6 on the Snellen visual acuity), non-demented cognitive status (≥21 on the Montreal cognitive assessment (MoCA) [12]), independently mobile indoors without a walking aid, absence of any neurological problem (other than PD for that group) or severe co-morbidity affecting gait.

PD specific inclusion criteria were; a diagnosis of idiopathic PD (by a consultant neurologist) and mild-moderately severe symptoms (Hoehn & Yahr (H&Y) stage I-III). PD participants were excluded if they presented with severe dyskinesia or experienced prolonged off periods. PD participants were tested on the peak dose of their medication.

2.2. Equipment

2.2.1. Mobile eye-tracker

A Dikablis (Ergoneers GmbH, Germany) mobile (head-mounted) infra-red eye-tracker measured saccade amplitude (distance between two fixations), which has an adequate sampling frequency (50 Hz) to detect saccades [11,13]. The Dikablis consisted of a light-weight head-unit and transmitter (weight: 69 g). The head-unit was double-sided taped to each participant's forehead to prevent slippage. The dual-camera system consisted of a monocular infra-red eye-camera to track pupil blackness and a fish-eye field-camera to record the environment in front of the participant. The system was calibrated using the manufacturer's four-point procedure (Fig. 1) for each participant. Calibration created a shared coordinate system relating the position of the pupil captured by the eye-camera with the gaze direction displayed on the field-of-view camera [11].

2.2.2. Head movement

Head and eye-movements are interdependent [14]. Head movement can impact saccade amplitude measurement when the head is unconstrained [15]. Therefore, head movement was recorded using a tri-axial accelerometer (Axivity AX3, York, 100 Hz) fixed to the Dikablis head-unit to examine whether head movement affected our findings.

2.3. Protocol

The study consisted of two sessions, one week apart. Accuracy was assessed using data from session 1 and re-test reliability was assessed using data from both sessions. Prior to testing, participants underwent demographic, clinical and cognitive assessments (MoCA and Mini Mental State Examination (MMSE)).

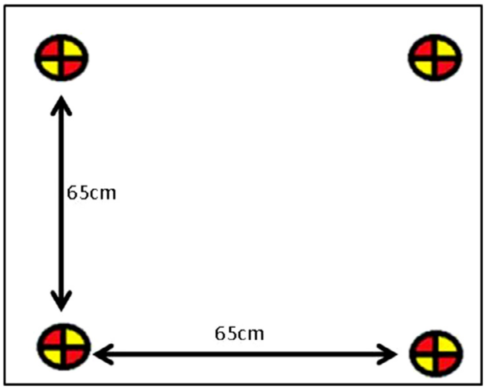

Fig. 1. Calibration board and procedure. Participants were seated and had a chin rest in place, and were then asked to move only their eyes to look at the targets on the board (65 cm square) starting at the bottom left target and continuing in a clockwise direction.

2.3.1. Accuracy (session 1)

Accuracy of saccade amplitude was examined by tracking eye-movements as participants looked between two targets placed at set distances (5°, 10° and 15°, Fig. 2) in time with a metronome (1 Hz) for 20 seconds (s). A maximal target distance of 15° was chosen as most naturally occurring saccades occur within this range [16]. Beyond 15°, co-ordinated eye-head movement is required [17]. Brief (30 s) rests were permitted after each trial to avoid fatigue, as previous studies have reported that fatigue occurs after a sequence of 36 s of eye-movements [18].

Eye-movement procedure:

Two highly salient targets (coloured red and yellow to attract visual attention) were placed on a white board 200 cm from the participant, with the central target at eye-level (Fig. 2). Participants were instructed to move their fixation from central to peripheral target (Fig. 2). Order was as follows:

(1) Horizontally: 5°,10°,15°
(2) Vertically: 5°,10°,15°

Tasks:

The eye-movement procedure was repeated during:

(1) Sitting (with chin rest; restricted head movement).
(2) Standing (asked not move their head; self-restricted head movement).
(3) Walking on a treadmill (Force Link, Netherlands) (head movement permitted). Treadmill speed was set to 80% of that achieved during a 10 m walk test carried out at the start of each session. Researchers provided verbal feedback to ensure participants stayed 2 m from the testing board.

2.3.2. Reliability

To assess re-test reliability, the same protocol described in Section 2.3.1 was repeated one week later (Mean: 7, SD: 2 days). All testing conditions were kept as consistent as possible, with trials conducted by the same researchers (SS, LA) using the same procedure, instructions and testing sequences.

2.3.3. Older adult without visual correction

To assess potential influence of visual correction (glasses or contact lenses) on accuracy and reliability, data from OA participants who did not require visual correction ($n = 10$) was re-analysed (Table 3).

Please cite this article as: S. Stuart et al., Accuracy and re-test reliability of mobile eye-tracking in Parkinson's disease and older adults, Medical Engineering and Physics (2016), http://dx.doi.org/10.1016/j.medengphy.2015.12.001

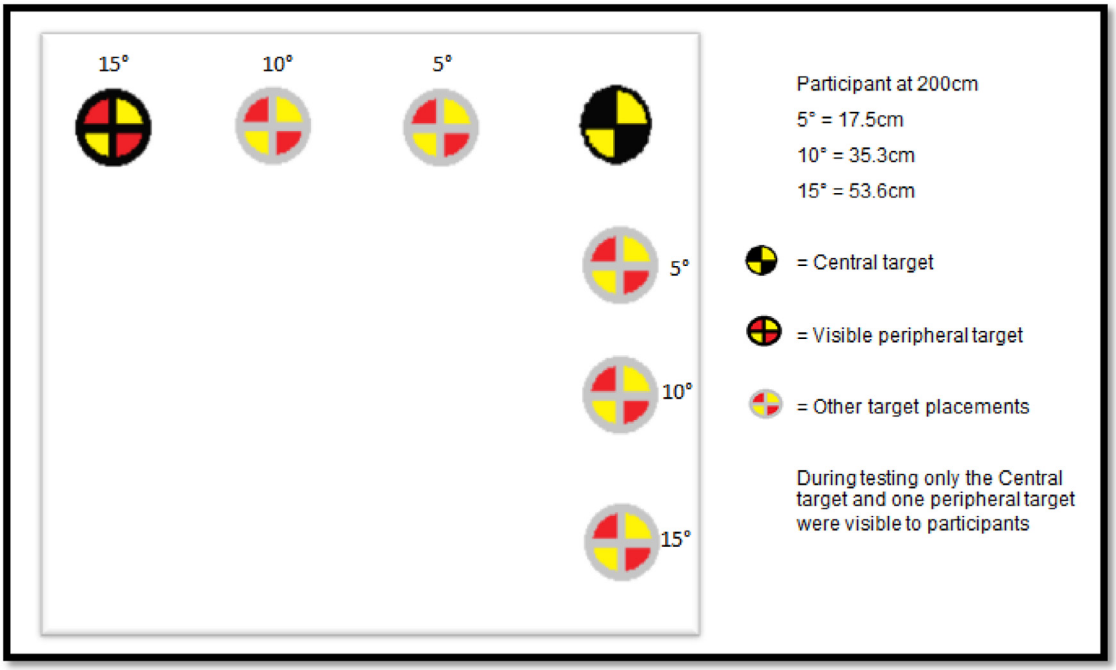

Fig. 2. Diagram illustrating the testing board used during sitting, standing and walking.

2.4. Data processing and analysis

2.4.1. Eye and head movement

Saccade amplitude and head movement were derived using a validated velocity-based algorithm (MATLAB® 2012a, Mathworks, USA) [11], which accounts for small ‘catch up’ saccades that follow large saccades to locate a target (i.e. saccades occurring within 100 ms of a previous saccade are summed to provide total distance). To quantify head movement impact on saccade amplitude, raw vertical and horizontal eye position data was compared to medio-lateral and superior-inferior head accelerations using cross-correlations (peak-correlation) as a measure of combined eye-head movement [19–22]. Head accelerations were low-pass filtered using a 4th order 30 Hz Butterworth filter [21,23].

2.4.2. Statistical analysis

Statistical analysis was performed using SPSS 21.0 (SPSS Inc., IL). Data were assessed for normality using Kolmogorov–Smirnov tests. Between groups comparison of saccade amplitude was not performed as this was not the study focus.

As the majority of variables were non-normally distributed, we did not calculate intra-class correlation. Instead, we describe accuracy in terms of bias and consistency of saccades. Bias was determined by subtracting known target distance from median saccade amplitude measured using the eye-tracker (median saccade amplitude—target distance). Consistency was calculated as the range (Maximum, Minimum) of error between measured and target saccade amplitude across participants.

Re-test reliability was described using median and range of between-session difference (median session 2—median session 1), and formally tested using Wilcoxon signed-rank tests for each target amplitude. Relative agreement between sessions was assessed using Spearman’s *rho* correlations. Correlation coefficients were interpreted as follows: excellent > 0.90, good ≥ 0.75–0.89, fair ≥ 0.50–0.74, and poor < 0.49 [24]. A threshold of $p<0.05$ guided interpretation.

Table 1
Demographics.

Characteristic	Older adults (n = 20) median (range)	Parkinson’s disease (n = 14) median (range)	p-value
Age (yrs),	68.5 (51, 86)	68.0 (61, 81)	.88
Sex, *n* (%)			
Men	12 (60%)	9 (64%)	.85
Women	8 (40%)	5 (36%)	
Height (cm)	170.5 (143, 184)	168.5 (150, 183)	.85
Weight (kg)	72.9 (58, 101)	78.3 (51, 107)	.36
Glasses, *n* (%)			
None	10 (50%)	2 (14.2%)	–
Bifocals	2 (10%)	4 (28.6%)	–
Varifocals	4 (20%)	4 (28.6%)	–
Contact lenses	3 (15%)	0 (0%)	–
Distance	1 (5%)	4 (28.6%)	–
Glasses worn during testing	10 (50%)	12 (86%)	.03∗
MMSE	30 (26, 30)	29 (24, 30)	.26
MoCA	28 (21, 30)	27 (23, 30)	.42
Years of education	13 (7, 20)	12 (10, 19)	.31
H & Y stage (*n*)	–	I (4), II (8), III (2)	–
UPDRS-III	–	34.5 (8, 63)	–
10 m walk (s)	7.73 (5.97, 13.84)	8.14 (6.01, 13.73)	.55
Walk speed (km/hr)	4.67 (2.61, 6.05)	4.43 (2.63, 6.01)	.58

MMSE: Mini-Mental State Examination; MoCA: Montreal Cognitive Assessment; UPDRS-III: Unified Parkinson’s disease Rating Scale – motor symptoms, H & Y stage: Hoehn and Yahr stage.
∗: $p<.05$.

3. Results

3.1. Demographics

Participant characteristics are described in Table 1. Some participants (OA $n = 2$, PD $n = 1$) were unable to complete session 2 but their data was retained for the accuracy analysis. There were no significant group differences in age, sex or education level. Participants wore their normal visual correction walk during walking,

Please cite this article as: S. Stuart et al., Accuracy and re-test reliability of mobile eye-tracking in Parkinson’s disease and older adults, Medical Engineering and Physics (2016), http://dx.doi.org/10.1016/j.medengphy.2015.12.001

with significantly more PD participants wearing visual correction (p = 0.03). The PD group had moderate motor symptoms as assessed using the MDS-UPDRS-III and H&Y-scale.

3.2. Eye and head movement

Low cross-correlation coefficients indicated that head movement did not influence saccade amplitude (r: 0.01–0.12 for walking, supplementary material 1). As such, standing and walking head movement data was not included in further analyses. The poor correlations were likely due to the maximum target distance of 15°, as saccades greater than 20° are needed to elicit combined eye-head movement [25,26].

3.3. Accuracy

Saccade amplitude consistently increased with target distance (Table 2). In relation to overall accuracy, bias of −1.23° and −1.17° was observed for PD and OA participants, respectively. However, poor consistency (large range of error between participants) was observed (PD: −7.48° to 5.18°; OA: −7.73° to 5.81°), which was dependent upon target distance (5°, 10°, 15°) and direction (horizontal, vertical). Task (sitting, standing, walking) did not significantly affect accuracy.

Table 2 shows that the magnitude of bias generally increased with the magnitude of eye-movement (e.g. sitting 5° = −0.19°, 10° = −2.69°, 15° = −5.66°). Similarly, both groups tended to 'undershoot' targets set 10° (e.g. −2.63°) and 15° (e.g. −4.94°) apart, which was consistent for all tasks. In addition, the range of error was greatest for larger saccades (e.g. 10° = −4.08° to 0.28° and 15° = −7.48° to 2.31°).

Bias was also related to saccade direction, such that participants undershot the target distance considerably more when performing vertical compared to horizontal saccades.

3.4. Reliability

Overall, median difference (session 2 – session 1) in saccade amplitude was low in both groups (PD; −0.14°, OA; 0.02°, Table 2). Similarly, median difference for individual tasks and amplitudes (Table 2) was low (<1°). Only one variable (OA; walking, horizontal, 15°) showed a significant difference between sessions (p = 0.02) but the median difference was still low (−0.95°). However, there was a wide range of difference between sessions across the participants (−12.60° to 16.75°). Relative agreement varied greatly from poor to good (*rho* range: 0.14, 0.85). Test condition did not have a consistent influence on bias or relative agreement. In contrast, larger saccades were associated with a greater range of change between sessions.

3.5. Influence of visual correction

Greater accuracy and reliability results were found in the subset of OA with no vision correction (Tables 2 and 3). With regards to accuracy, median bias from target reduced from −1.17° to −1.15° and error was more consistent across the participants. Median difference in saccadic amplitude between sessions (reliability) was similar but between-person range was much smaller. Modest improvements were also seen in relative agreement between sessions when considering people who did not use visual correction.

4. Discussion

To our knowledge, this is the first study to examine accuracy and reliability of a mobile eye-tracker in people with PD and OA. Results provide evidence that mobile eye-trackers can measure saccade amplitude in people with PD and OA although the accuracy and reliability depend on several factors. Findings contribute to the development of novel protocols for establishing the psychometric properties of mobile eye-trackers.

4.1. Accuracy

Median saccade amplitude measured by the mobile eye-tracker, increased with increasing target distance (Table 2). This indicates that the mobile eye-tracker can discern change in saccade amplitude. However, the measured saccade amplitudes were smaller than target distance, especially for larger and vertical saccades. In addition, bias was inconsistent across the participants, especially for larger saccades.

Although our previous work has shown mobile eye-trackers can accurately detect saccade occurrence [11], this study indicates saccade amplitude may not be measured with the same degree of certainty. This suggests that saccade detection outcomes (number/frequency) are more robust than saccade amplitude. Regardless, overall median bias (−1.21°) and consistency (−7.73° to 5.81°) is acceptable for certain protocols, such as dynamic protocols involving saccade detection which often use a minimum threshold of ≥5° saccade amplitude [2] to account for artefact error (e.g. vestibular-ocular-reflex) [11]. However, this degree of accuracy may not be acceptable for protocols where precision of large saccade amplitude is important.

4.2. Reliability

Re-test reliability varied across conditions and participants. Although median difference between sessions was low (<1°), difference ranged from −12.60° to 16.75° across participants. Similarly, relative agreement ranged from poor to good between conditions (*rho*; 0.14–0.85). Variable reliability indicates that saccade amplitude measurement may not be stable over time and is likely due to several sources of error (Section 4.3). Until robust protocols are developed which are stable over time, we cannot recommend saccade amplitude as a reliable mobile eye-tracker outcome.

4.3. Challenges and recommendations

Error noted in both accuracy and reliability stems from technological, human and study-protocol factors. A better understanding of these sources of error is important for future protocols and devices.

4.3.1. Technology factors

Manufacturer reported accuracy (0.5°) was not observed in this study. In contrast, a preliminary study (four young adults) using a static eye-tracker (Tobii, TX300; 300 Hz) during treadmill walking reported eye-tracker accuracy was consistent with manufacturer specifications (0.5°) regardless of target locations or saccade amplitude [9,27]. Inconsistency between the current study and the previous preliminary work [9], may be due to the lower sampling-frequency of the mobile eye-tracker used in this study (50 Hz) compared to the static device (300 Hz) [10]. A sampling-frequency of 50 Hz enables saccade detection [13] but a higher frequency (>200 Hz) may be more accurate at reporting specific characteristics [1]. A sampling-frequency of 50 Hz assumes that the eye is in a fixed location for 20 ms (50 Hz) whereas a higher frequency system (1000 Hz) assumes this for only 1 ms, providing better temporal accuracy and more eye-position data [10,13]. Therefore, a mobility-accuracy trade-off exists. Static higher sampling-frequency devices may offer improved accuracy and reliability but for pragmatic use, studies must limit participant mobility during dynamic tasks. That is, participants must walk on a treadmill and

Please cite this article as: S. Stuart et al., Accuracy and re-test reliability of mobile eye-tracking in Parkinson's disease and older adults, Medical Engineering and Physics (2016), http://dx.doi.org/10.1016/j.medengphy.2015.12.001

Table 2
Accuracy (session 1) and re-test reliability (comparison between session 1 and session 2).

	Task	Direction	Accuracy (session 1) (saccade amplitude (°))				Re-test reliability (session 2) (saccade amplitude (°))				
			°	Median (Min, Max)	Bias	Range of error	Median (Min, Max)	Median difference	Range of difference	*p*-value	Spearman's rho (*p*-value)
Older adults (*n=20*)	**Sitting**	**Horizontal**	**5**	5.69 (4.84, 9.56)	0.69	−0.16, 4.56	5.96 (4.41, 8.08)	−0.03	−5.51, 2.20	0.98	0.42 (0.07)
			10	10.23 (7.66, 13.18)	0.23	−2.34, 3.18	9.87 (8.59, 13.50)	−0.09	−8.28, 3.35	0.60	0.35 (0.14)
			15	12.71 (9.87, 14.52)	−2.29	−0.13, 4.52	13.28 (10.93, 14.71)	0.45	−11.76, 2.03	0.27	0.20 (0.42)
		Vertical	**5**	4.88 (4.05, 7.00)	−0.12	−0.95, 2.00	5.13 (4.05, 21.09)	0.21	−7.00, 16.75	0.14	0.34 (0.16)
			10	7.42 (6.20, 11.77)	−2.58	−3.80, 1.77	7.74 (6.34, 20.90)	0.07	−6.52, 12.53	0.32	0.27 (0.27)
			15	9.55 (7.27, 13.70)	−5.45	−7.73, −1.30	9.84 (7.85, 20.70)	0.26	−8.37, 12.15	0.29	0.27 (0.38)
	Median			–	−1.21	−7.73, 4.56	–	–	–	–	–
	Standing	**Horizontal**	**5**	6.16 (4.77, 10.81)	1.16	−0.23, 5.81	6.38 (4.98, 9.76)	−0.22	−6.23, 4.64	0.90	0.48 (0.30)
			10	10.01 (4.77, 10.81)	0.01	−5.23, 4.77	10.57 (8.48, 14.46)	0.39	−7.92, 2.62	0.55	0.36 (0.13)
			15	12.68 (10.51, 14.77)	−2.32	−4.49, −0.23	13.22 (10.91, 13.99)	0.06	−11.69, 2.83	0.81	0.21 (0.39)
		Vertical	**5**	5.15 (3.98, 10.38)	0.15	−1.02, 5.38	4.98 (4.05, 15.96)	−0.27	−4.65, 11.13	0.35	0.30 (0.21)
			10	7.55 (5.81, 11.97)	−2.45	−4.19, 1.97	7.58 (5.95, 19.03)	0.32	−6.22, 11.32	0.11	0.61 (0.005)
			15	10.17 (7.96, 12.00)	−4.83	−7.04, −3.00	9.79 (7.11, 21.15)	−0.36	−8.68, 9.16	0.89	0.66 (0.002)
	Median			–	−1.16	−7.04, 5.81	–	–	–	–	–
	Walking	**Horizontal**	**5**	5.41 (4.68, 8.16)	0.41	−0.32, 3.16	5.81 (4.30, 9.60)	0.21	−5.59, 4.92	0.07	0.30 (0.28)
			10	9.59 (7.02, 14.48)	−0.41	−2.98, 4.48	9.44 (7.33, 13.79)	−0.55	−8.71, 3.05	0.88	0.26 (0.29)
			15	13.07 (9.55, 14.37)	−1.93	−5.45, −0.63	11.96 (10.25, 13.41)	−0.95	−12.60, 3.51	0.02*	0.14 (0.57)
		Vertical	**5**	4.93 (4.46, 7.24)	−0.07	−0.54, 2.24	5.22 (4.17, 7.53)	−0.04	4.90, 2.97	0.34	0.53 (0.24)
			10	7.22 (5.52, 9.35)	−2.78	−4.28, −0.65	7.43 (5.86, 9.12)	−0.09	−6.67, 2.10	1.00	0.45 (0.06)
			15	10.21 (7.87, 12.01)	−4.79	−7.13, −2.99	10.63 (7.93, 12.06)	0.10	−8.22, 2.86	0.32	0.75 (0.001)
	Median			–	−1.17	−7.13, 4.48	–	–	–	–	–
	Group median			–	−1.17	−7.73, 5.81	–	0.02	−12.60, 12.53	–	–
Parkinson's disease (*n=14*)	**Sitting**	**Horizontal**	**5**	5.81 (4.45, 6.74)	0.81	−0.55, 1.74	6.10 (4.99, 7.74)	0.05	−5.18, 3.19	0.27	0.17 (0.59)
			10	9.52 (7.02, 13.40)	−0.48	−2.98, 3.40	9.80 (7.59, 12.69)	−0.25	−9.08, 2.88	0.89	0.51 (0.07)
			15	12.31 (8.80, 14.98)	−2.69	−6.20, −0.02	12.56 (10.24, 14.01)	−0.02	−11.40, 2.42	0.91	0.37 (0.29)
		Vertical	**5**	4.81 (4.03, 6.26)	−0.19	−0.97, 1.26	4.76 (4.05, 6.87)	−0.29	−4.51, 2.12	0.36	0.14 (0.65)
			10	7.31 (6.01, 9.00)	−2.69	−3.99, −1.00	7.00 (6.04, 10.84)	−0.55	−6.97, 2.62	0.69	0.64 (0.18)
			15	9.34 (7.80, 11.70)	−5.66	−7.20, −3.30	9.25 (7.89, 11.19)	−0.31	−8.65, 1.23	0.46	0.67 (0.01)
	Median			–	−1.59	−7.20, 3.40	–	–	–	–	–
	Standing	**Horizontal**	**5**	5.94 (4.81, 10.18)	0.94	−0.19, 5.18	6.05 (4.32, 7.59)	−0.13	−5.32, 1.37	0.73	0.76 (0.002)
			10	10.13 (8.20, 12.08)	0.13	−1.80, 2.08	10.28 (6.91, 13.50)	−0.21	−9.53, 2.23	0.24	0.85 (0.000)
			15	12.20 (9.90, 13.62)	−2.80	−5.10, −1.38	12.50 (10.13, 17.47)	0.45	−10.63, 5.03	0.15	0.64 (0.02)
		Vertical	**5**	4.79 (4.25, 5.53)	−0.21	−0.75, 0.53	4.56 (3.91, 11.08)	−0.08	−4.58, 6.63	0.37	0.38 (0.20)
			10	8.02 (6.10, 12.25)	−1.98	−3.90, 2.25	7.52 (6.08, 10.14)	−0.41	−6.63, 1.42	0.51	0.38 (0.20)
			15	9.82 (7.54, 11.91)	−5.18	−7.46, −3.09	9.11 (7.19, 12.54)	−0.75	−8.65, 1.10	0.10	0.50 (0.08)
	Median			–	−1.10	−7.46, 5.18	–	–	–	–	–
	Walking	**Horizontal**	**5**	5.62 (4.65, 9.90)	0.62	−0.35, 4.90	5.58 (4.95, 6.24)	−0.01	−5.15, 0.91	0.62	0.20 (0.51)
			10	9.70 (6.29, 12.94)	−0.30	−3.71, 2.94	9.93 (7.99, 13.00)	0.15	−8.82, 2.11	0.20	0.63 (0.02)
			15	12.38 (8.53, 13.82)	−2.62	−6.47, −1.18	12.92 (11.09, 15.67)	0.23	−11.40, 5.24	0.16	0.14 (0.65)
		Vertical	**5**	4.80 (4.35, 6.98)	−0.20	−0.65, 1.98	4.68 (4.32, 5.77)	−0.15	−4.45, 0.72	0.10	0.44 (0.13)
			10	7.37 (5.92, 10.28)	−2.63	−4.08, 0.28	6.95 (5.83, 16.30)	−0.11	−6.63, 6.55	0.67	0.45 (0.13)
			15	10.06 (7.52, 12.31)	−4.94	−7.48, 2.31	9.52 (7.28, 11.67)	−0.27	−8.68, 1.45	0.21	0.80 (0.001)
	Median			–	−1.46	−7.48, 4.90	–	–	–	–	–
	Group median			–	−1.23	−7.48, 5.18	–	−0.14	−11.40, 5.24	–	–
	Overall median			–	−1.21	−7.73, 5.81	–	−0.09	−12.60, 16.75	–	–

*Significance level $p<0.05$.

Table 3
Accuracy (Session 1) and re-test reliability (comparison of Session 1 and Session 2) of older adults with no vision correction ($n = 10$).

		Accuracy (saccade amplitude (°))				Re-test reliability (saccade amplitude (°))				
		session 1				session 2				
Task	Direction	°	Median (Min, Max)	Bias	Range of error	Median (Min, Max)	Median difference	Range of difference	*p*-value	Spearman's rho (*p*-value)
Sitting	**Horizontal**	**5**	5.58 (4.84, 7.48)	0.58	−0.16, 2.48	5.91 (5.21, 6.98)	0.24	−0.52, 1.34	0.14	0.29 (0.42)
		10	9.86 (7.66, 12.35)	−0.14	−2.34, 2.35	9.48 (8.59, 13.50)	−0.09	−2.87, 3.35	1.00	0.89 (0.05)
		15	13.13 (9.87, 14.52)	−1.87	−5.13, −0.48	12.78 (10.93, 14.54)	0.27	−2.10, 1.63	0.95	0.33 (0.35)
	Vertical	**5**	4.75 (4.05, 5.35)	−0.25	−0.95, 0.35	4.88 (4.05, 5.42)	0.04	−0.83, 0.94	0.36	0.13 (0.73)
		10	6.76 (6.20, 9.03)	−3.24	−3.80, −0.97	7.42 (6.40, 9.00)	0.43	−2.30, 1.78	0.26	0.83 (0.08)
		15	9.14 (7.27, 10.88)	−5.86	−7.73, −4.12	9.70 (7.85, 11.44)	0.64	−1.04, 1.43	0.07	0.76 (0.01)
	Median		–	−1.06	−7.73, 2.48	–	–	–	–	–
Standing	**Horizontal**	**5**	5.97 (4.77, 7.17)	0.97	−0.23, 2.17	5.89 (4.98, 7.47)	0.23	−0.56, 1.44	0.38	0.77 (0.009)
		10	10.01 (7.98, 14.42)	0.01	−2.02, 4.42	10.41 (8.48, 12.61)	0.20	−2.59, 2.62	0.84	0.32 (0.36)
		15	12.80 (10.85, 14.77)	−2.20	−4.15, 4.77	13.20 (10.91, 13.84)	−0.06	−1.42, 1.96	0.92	0.20 (0.59)
	Vertical	**5**	4.76 (3.98, 6.10)	−0.24	−1.02, 1.10	4.92 (4.05, 5.57)	0.12	−1.06, 1.18	0.88	0.17 (0.65)
		10	6.57 (5.81, 8.16)	−3.43	−4.19, −1.84	7.04 (5.95, 8.32)	0.32	−1.27, 1.61	0.26	0.53 (0.12)
		15	9.55 (7.96, 11.12)	−5.45	−7.04, −3.88	8.82 (7.11, 10.43)	−0.48	−2.89, 0.70	0.15	0.43 (0.21)
	Median		–	−1.22	−7.04, 4.77	–	–	–	–	–
Walking	**Horizontal**	**5**	5.40 (4.80, 5.77)	0.40	−0.20, 0.77	5.76 (4.30, 6.13)	0.09	−4.80, 0.82	0.37	0.40 (0.28)
		10	9.93 (7.02, 14.30)	−0.07	−2.98, 4.30	8.86 (7.33, 13.23)	−0.63	−8.37, 2.30	0.40	0.23 (0.56)
		15	13.85 (10.46, 14.37)	−1.15	−4.56, 4.37	12.47 (10.82, 13.41)	−1.19	−10.49, 0.12	0.008∗	0.43 (0.25)
	Vertical	**5**	4.81 (4.58, 7.24)	−0.19	−0.42, 2.24	5.24 (4.17, 6.11)	0.19	−4.90, 0.72	0.40	0.44 (0.24)
		10	7.14 (4.58, 7.24)	−2.86	−5.42, −2.76	6.83 (5.86, 8.05)	−0.09	−6.29, 0.95	0.35	0.42 (0.27)
		15	9.97 (7.87, 10.89)	−5.03	−7.13, −4.11	9.21 (7.93, 11.08)	0.04	−8.01, 0.84	1.00	0.74 (0.02)
	Median		–	−0.67	−7.13, 4.37	–	–	–	–	–
	Overall median		–	−1.15	−7.73, 4.77	–	0.11	−10.49, 3.35	–	–

$*p<0.05$.

be at a set distance from visual targets [9]. However, protocols which limit mobility impact validity of the characteristics measured [28], e.g., restricted head movements during static protocols may facilitate abnormal visual processing, seen through alterations in saccade responses [29].

Some bias may be due to eye curvature induced error [30]. The eye is a convex curved lens with a horizontal movement range of ∼100° and vertical range of ∼90° [31]. Many eye-trackers locate the pupil via the black pixels recorded by an infra-red eye-camera and uses specific circular pupil shape parameters to derive the pupil centre. Depending upon the location of the eye-camera in relation to the eye, the pupil shape will appear as an ellipse and therefore the circular pupil shape parameters would lead to inaccurate tracking. This is relevant for large saccades, where the person is looking furthest from the camera. The Dikablis eye-tracker demonstrated such an error by recording an 'undershoot' for all targets at 15° and may have contributed to the poorer accuracy for 15° saccades. This error could be controlled for in future technology with the use of convex cost function algorithms [32] or corneal reflexion tracking [33], which would provide further means of tracking eye-in-head movements [34] and control for pupil tracking errors [35].

4.3.2. *Visual correction and obstruction of the eye*

Pupil tracking was likely impacted by a number of general eye-tracker issues, such as inaccuracies due to poor calibration [36], long or drooping eye lashes/lids, infra-red refraction due to visual correction (glasses), hair obstruction and any slippage of the 'one-size-fits-all' eye-tracker from original placement when recording [13]. During data collection eye lids/lashes and visual correction (particularly bi-focal glasses) were observed as main cause of error, particularly for vertical saccades and large saccades of any direction. These challenges are inherent to infra-red eye-tracking devices and although some can be controlled within an experiment, many are dependent upon researcher ability to identify and address these issues. Using double-sided tape to minimise device slippage and requesting participants not wear eye make-up were ways which we found anecdotally improved accuracy.

We assessed whether visual correction may have impacted accuracy and re-test reliability by looking at a subset of 10 OA who wore no visual correction. Results showed that the accuracy and reliability were better in individuals who did not use visual correction, likely due to visual correction affecting pupil detection via infra-red refraction [13]. Unfortunately, exclusion of participants with visual correction may not be appropriate when selecting participants, particularly with groups likely to have increased use of visual correction such as OA. Therefore, the negative effect of visual correction on eye-tracker accuracy and reliability must be considered when designing robust protocols and is a challenge which still needs to be addressed by manufacturers of the next generation of eye-trackers.

4.3.3. *Visual attention*

Participant saccades were voluntary and therefore involved selective visual attention which is influenced by internal factors [37] and may have affected amplitude results. Factors such as level of fatigue between sessions [38], ethnicity of participants [39], prior knowledge of testing protocols (learning effect) [40], individual emotional state [41] and motivation [42] could all have influenced saccade measures. Future studies could control for such factors by investigating saccade latencies compared to auditory signal, or quantifying total saccade number to compare to a set amount (i.e. 20 saccades within 20 seconds).

In addition, this study did not consider the inhibition-of-return mechanism whereby a person orientates their attention to novel locations and stimuli, as our target appearance, location and saliency [43] remained the same. Once a peripheral location is foveated (fixated) there is a delayed response in returning attention to subsequent stimuli in the same location [44]. Programming of the next saccade occurs even before the previous saccade is completed [45], therefore introducing a time constraint (1 s) and using the same targets/locations may have led to inaccuracies in saccade programming and execution. Therefore, some of the error observed in this study may have been due to inaccurate saccades rather than error introduced by the mobile eye-tracker.

Please cite this article as: S. Stuart et al., Accuracy and re-test reliability of mobile eye-tracking in Parkinson's disease and older adults, Medical Engineering and Physics (2016), http://dx.doi.org/10.1016/j.medengphy.2015.12.001

4.3.4. Limitations

Future work should address the limitations of this study to establish a ‘gold standard’ method to be applied to differing devices and various populations. Novel peripheral targets in varying locations which require reflexive (involuntary) saccades should be used, with variations on saccadic timings. For example; a light board or computer-based programme where objects or targets randomly appear (similar to that used by Serchi, Peruzzi [9] for their static eye-tracker) could be used with mobile devices. Future studies could also examine the impact of combined eye-head movement on saccade amplitude accuracy, particularly for larger saccades (>20°) where coordinated eye-head movement is required.

5. Conclusion

The Dikablis mobile eye-tracker had variable accuracy and reliability when recording saccade amplitude in people with PD and OA. Accuracy is acceptable for certain protocols but more precision may be necessary when investigating specific saccade characteristics. Error was induced via several technological, human and study-specific factors which need to be addressed to achieve more robust testing protocols.

Conflict of interest

None.

Funding

This research is supported by the National Institute for Health Research (NIHR) Newcastle Biomedical Research Unit/Centre (BRU/BRC) based at Newcastle-upon-Tyne Hospitals NHS Foundation Trust and Newcastle University. The research was also supported by NIHR Newcastle CRF Infrastructure funding. The views expressed are solely those of the authors.

Ethical approval

Ethical approval was obtained via local research ethics committee (Newcastle and North-Tyneside REC-1; 13/NE/0128). Participants provided written informed consent prior to testing.

Acknowledgements

The authors acknowledge Aodhán Hickey (Newcastle University) for his technical support.

Supplementary materials. Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.medengphy.2015.12.001.

References

[1] Stuart S, Alcock L, Galna B, Lord S, Rochester L. The measurement of visual sampling during real-world activity in Parkinson’s disease and healthy controls: A structured literature review. J Neurosci Methods 2014;222:175–88.
[2] Galna B, Lord S, Daud D, Archibald N, Burn D, Rochester L. Visual sampling during walking in people with Parkinson’s disease and the influence of environment and dual-task. Brain Res 2012;1473:35–43.
[3] Pelz JB, Canosa R. Oculomotor behavior and perceptual strategies in complex tasks. Vis Res 2001;41:3587–96.
[4] Blekher T, Weaver MR, Cai X, Hui S, Marshall J, Jackson JG, et al. Test-retest reliability of saccadic measures in subjects at risk for Huntington disease. Invest Ophthalmol Vis Sci 2009;50:5707–11.
[5] Klein C, Fischer B. Instrumental and test-retest reliability of saccadic measures. Biol Psychol 2005;68:201–13.
[6] Farris-Trimble A, McMurray B. Test-retest reliability of eye tracking in the visual world paradigm for the study of real-time spoken word recognition. J Speech, Lang, Hear Res: JSLHR 2013;56:1328–45.
[7] Farzin F, Scaggs F, Hervey C, Berry-Kravis E, Hessl D. Reliability of eye tracking and pupillometry measures in individuals with fragile X syndrome. J Autism Dev Disord 2011;41:1515–22.
[8] Ettinger U, Kumari V, Crawford TJ, Davis RE, Sharma T, Corr PJ. Reliability of smooth pursuit, fixation, and saccadic eye movements. Psychophysiology 2003;40:620–8.
[9] Serchi V, Peruzzi A, Cereatti A, Della Croce U. Tracking gaze while walking on a treadmill: spatial accuracy and limits of use of a stationary remote eye-tracker. In: 36th annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2014; 2014. p. 3727–30.
[10] Andersson R, Nyström M, Holmqvist K. Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. J Eye Mov Res 2010;3:1–12.
[11] Stuart S, Galna B, Lord S, Rochester L, Godfrey A. Quantifying Saccades while walking: validity of a novel velocity-based algorithm for mobile eye tracking. In: 36th annual international conference of the IEEE engineering in medicine and biology society, 2014 EMBS 2014, Chicago, Illinois, USA. IEEE; 2014.
[12] Bernardi G, Mild M. Cognitive Impairment is Under-recognised in newly referred patients with parkinson’s disease: mini mental state examination (mmse) versus Montreal cognitive assessment (Moca). J Alzheimer’s Dis Park 2012:01.
[13] Holmqvist K, Nystrom M. Eye tracking: a comprehensive guide to methods and measures. Oxford, UK: Oxford University Press; 2011.
[14] Freedman EG. Interactions between eye and head control signals can account for movement kinematics. Biol Cybern 2001;84:453–62.
[15] Proudlock FA, Shekhar H, Gottlob I. Age-related changes in head and eye coordination. Neurobiol Aging 2004;25:1377–85.
[16] Bahill AT, Adler D, Stark L. Most naturally occurring human saccades have magnitudes of 15 degrees or less. Invest Opthalmol 1975;14:468–9.
[17] Maurer C, Mergner T, Lucking CH, Becker W. Adaptive changes of saccadic eye–head coordination resulting from altered head posture in torticollis spasmodicus. Brain 2001;124:413–26.
[18] Wilson SJ, Glue P, Ball D, Nutt DJ. Saccadic eye movement parameters in normal subjects. Electroencephalograph Clin Neurophysiol 1992;86:69–74.
[19] Lee C. Eye and head coordination in reading: roles of head movement and cognitive control. Vis Res 1999;39:3761–8.
[20] Pelz J, Hayhoe M, Loeber R. The coordination of eye, head, and hand movements in a natural task. Exp Brain Res 2001;139:266–77.
[21] Kavanagh JJ, Barrett RS, Morrison S. Upper body accelerations during walking in healthy young and elderly men. Gait Posture 2004;20:291–8.
[22] Kavanagh JJ, Menz HB. Accelerometry: a technique for quantifying movement patterns during walking. Gait Posture 2008;28:1–15.
[23] Kavanagh JJ, Morrison S, Barrett RS. Coordination of head and trunk accelerations during walking. Eur J Appl Physiol 2005;94:468–75.
[24] Rosner BA. Fundamentals of biostatistics 2006.
[25] Crawford JD, Martinez-Trujillo JC, Klier EM. Neural control of three-dimensional eye and head movements. Curr Opin Neurobiol 2003;13:655–62.
[26] Gandhi NJ, Sparks DL. Experimental control of eye and head positions prior to head-unrestrained gaze shifts in monkey. Vis Res 2001;41:3243–54.
[27] Serchi V, Peruzzi V, Cereatti A, Della Croce U. Performance of a remote eye-tracker in measuring gaze during walking. In: 20th IMEKO TC4 International Symposium and 18th International Workshop on ADC Modelling and Testing Research on electric and electronic measurement for the economic upturn; 2014. p. 770–4.
[28] Nevalainen S, Sajaniemi J. Comparison of three eye tracking devices in psychology of programming research. In: Proceedings of the 16th annual workshop of the psychology of programming interest group; 2004. p. 151–8.
[29] van Stockum S, MacAskill MR, Myall D, Anderson TJ. A perceptual discrimination task results in greater facilitation of voluntary saccades in Parkinson’s disease patients. Eur J Neurosci 2013;37:163–72.
[30] Zhiwei Z, Qiang J. Novel eye gaze tracking techniques under natural head movement. IEEE Trans Biomed Eng 2007;54:2246–60.
[31] Botha CP, de Graaf T, Schutte S, Root R, Wielopolski P, van der Helm F, L Linsen, et al. MRI-based Visualisations of orbital fat deformation during eye motion Visualization in medicine and life sciences. Hagen H, Hamann B, editors. Berlin Heidelberg: Springer-Verlag; 2008.
[32] De Santis A, Iacoviello D. Robust real time eye tracking for computer interface for disabled people. Comput Methods Programs Biomed 2009;96:1–11.
[33] Mele ML, Federici S. A psychotechnological review on eye-tracking systems: towards user experience. Disabil Rehabil Assist Technol 2012;7:261–81.
[34] Hennessey CA, Lawrence PD. Improving the accuracy and reliability of remote system-calibration-free eye-gaze tracking. IEEE Trans Bio-med Eng 2009;56:1891–900.
[35] Li F, Munn S, Pelz J. A model-based approach to video-based eye tracking. J Mod Opt 2008;55:503–31.
[36] Nystrom M, Andersson R, Holmqvist K, van de Weijer J. The influence of calibration method and eye physiology on eyetracking data quality. Behav Res Methods 2013;45:272–88.
[37] Baluch F, Itti L. Mechanisms of top-down attention. Trends Neurosci 2011;34:210–24.
[38] Faber LG, Maurits NM, Lorist MM. Mental fatigue affects visual selective attention. PloS One 2012;7:e48073.

Please cite this article as: S. Stuart et al., Accuracy and re-test reliability of mobile eye-tracking in Parkinson’s disease and older adults, Medical Engineering and Physics (2016), http://dx.doi.org/10.1016/j.medengphy.2015.12.001

[39] Blignaut P, Wium D. Eye-tracking data quality as affected by ethnicity and experimental design. Behav Res Methods 2014;46:67–80.
[40] Kim S, Rehder B. How prior knowledge affects selective attention during category learning: an eyetracking study. Mem Cogn 2011;39:649–65.
[41] Oatley K, Parrott WG, Smith C, Watts F. Cognition and emotion over twenty-five years. Cognit Emot 2011;25:1341–8.
[42] Kaplan RL, Van Damme I, Levine LJ. Motivation matters: differing effects of pre-goal and post-goal emotions on attention and memory. Front Psychol 2012;3:404.
[43] t Hart B.M., Schmidt H.C.E.F., Klein-Harmeyer I., Einhäuser W. Attention in natural scenes: contrast affects rapid visual processing and fixations alike 2013.
[44] Klein RM. Inhibition of return. Trends Cognit Sci 2000;4:138–47.
[45] McPeek RM, Skavenski AA, Nakayama K. Concurrent processing of saccades in visual search. Vis Res 2000;40:2499–516.

Please cite this article as: S. Stuart et al., Accuracy and re-test reliability of mobile eye-tracking in Parkinson's disease and older adults, Medical Engineering and Physics (2016), http://dx.doi.org/10.1016/j.medengphy.2015.12.001