Is it reliable to assess visual attention of drivers affected by Parkinson’s disease from the backseat?—a simulator study

Hoe C. Lee¹*, Derserri Yanting Chee¹, Helena Selander²,³ and Torbjorn Falkmer¹,²,⁴,⁵

¹School of Occupational Therapy and Social Work, Curtin Health Innovation Research Institute, Curtin University, Perth, WA, Australia; ²School of Health Sciences, Jönköping University, Jönköping, Sweden; ³Mobilitetscenter, Gothenburg, Sweden; ⁴Department of Rehabilitation Medicine, IKE, Faculty of Health Sciences, Linköping University, Linköping, Sweden; ⁵School of Occupational Therapy, La Trobe University, Melbourne, VIC, Australia

Background: Current methods of determining licence retainment or cancellation is through on-road driving tests. Previous research has shown that occupational therapists frequently assess drivers’ visual attention while sitting in the back seat on the opposite side of the driver. Since the eyes of the driver are not always visible, assessment by eye contact becomes problematic. Such procedural drawbacks may challenge validity and reliability of the visual attention assessments. In terms of correctly classified attention, the aim of the study was to establish the accuracy and the inter-rater reliability of driving assessments of visual attention from the back seat. Furthermore, by establishing eye contact between the assessor and the driver through an additional mirror on the wind screen, the present study aimed to establish how much such an intervention would enhance the accuracy of the visual attention assessment.

Methods: Two drivers with Parkinson’s disease (PD) and six control drivers drove a fixed route in a driving simulator while wearing a head mounted eye tracker. The eye tracker data showed where the foveal visual attention actually was directed. These data were time stamped and compared with the simultaneous manual scoring of the visual attention of the drivers. In four of the drivers, one with Parkinson’s disease, a mirror on the windscreen was set up to arrange for eye contact between the driver and the assessor. Inter-rater reliability was performed with one of the Parkinson drivers driving, but without the mirror.

Results: Without mirror, the overall accuracy was 56% when assessing the three control drivers and with mirror 83%. However, for the PD driver without mirror the accuracy was 94%, whereas for the PD driver with a mirror the accuracy was 90%. With respect to the inter-rater reliability, a 73% agreement was found.

Conclusion: If the final outcome of a driving assessment is dependent on the subcategory of a protocol assessing visual attention, we suggest the use of an additional mirror to establish eye contact between the assessor and the driver. The clinicians’ observations on-road should not be a standalone assessment in driving assessments. Instead, eye trackers should be employed for further analyses and correlation in cases where there is doubt about a driver’s attention.

Keywords: cognitive deficits; driver assessment; eye contact; eye tracking; fixations; foveation; occupational therapy; rear view mirror

Received: 12 December 2011; Revised: 29 January 2012; Accepted: 31 January 2012; Published: 27 February 2012

Parkinson’s Disease (PD) is the second most common neurological disease in Australia; causing impairments in motor control, cognitive functioning and sensation (1). The neurodegenerative disease can impair functional driving performance and increase the risk of crashes and fatalities on Australian roads (2). In particular, cognitive symptoms of PD can have a substantial influence on driving performance, due to the complicated and demanding nature of the task (3). Research into the impact of cognitive symptoms upon driving ability is limited and contradictory. It is difficult to detect the presence of cognitive impairment in PD and to determine the relationship and severity of cognitive impairment on driving performance. The exact
prevalence of cognitive impairment amongst drivers with PD is unknown and difficult to ascertain (4).

In Australia, the guidelines regulating licence retention and cancellation of drivers affected by Parkinson’s Disease (PD) are based upon a system of subjective medical expert opinion (5). There are no current national standards or requirements for how clinical driving assessments should be conducted (6). Specific clinical assessment batteries and criteria to renew or cancel driving licences have not been clearly defined (7). The combination of symptoms and/or the severity that could compromise driving ability are not defined. Therefore, the medical practitioner must make a subjective decision on the fitness to drive of their patients, even though they may not have been trained to do so (7).

The cheapest, most accessible and commonly used method for determining driving ability is through clinical assessment. Tools, such as the Timed Up and Go (measures ability to stand up, walk for three metres and return to the chair), Unified Parkinson’s Scale and Mini Mental Status Examination (MMSE) are commonly used (7). However, the predictive validity of using these tools in driving assessment of PD drivers is frequently questioned in the literature (7–10). Ernst and Paulus 2005 noted that it is difficult to assess risk-taking behaviours in an indoor, clinical setting without actually watching the person drive (11). In a double blind study using 20 people with PD and 20 age-matched controls; it was found that there was a 35% inconsistency in clinical assessment results conducted by a neurologist, compared to on-road driving assessment results provided by a driving instructor and occupational therapist (4). Although these results need to be interpreted with caution due to the small sample size; it does highlight that assessment processes need to be improved. Betz and Fisher 2009 suggested that further research into the detection of cognitive impairment and its potential implications for road safety is becoming more crucial in preventing fatal collisions as the population ages (9).

However, Heikkila et al. did suggest that visual memory, choice reaction time and information processing speed tests could potentially be used to assess fitness to drive; once more research is conducted to establish validity and reliability (4). Moreover, the Heikkila et al. study suggested that visual memory, choice reaction time and information processing speed tests could potentially be used to assess fitness to drive; once more research is conducted to establish validity and reliability. Previous research has highlighted eye-trackers can accurately measure the variables (12–17), however, this has not yet been researched in relation to people with PD. Current methods of determining licence retention or cancellation is through on-road driving tests and/or clinical psychometric assessments (18). On-road assessment is the gold standard. However, the process is costly and time consuming (19, 20). It is therefore the responsibility of the occupational therapy profession to continue to develop knowledge in the area of driving, to improve road safety and support Occupational Therapists (OTs) working in the field (21). This will also assist in fulfilling legal, social and professional responsibilities and enable occupational therapists to justify their role in working with drivers with PD (20, 22).

Previous research has shown that OTs frequently assess drivers’ behaviour while sitting on the opposite side of the driver in the back seat (23, 24). Commonly, driving assessments protocols have subtasks assessing attention, for example the valid and reliable P-Drive (24), where 5 of the 27 subtasks (19%) are directly addressing ‘attention’. They are attending straight ahead; to the right; to the left; to mirrors and to fellow road users. Another example of protocols addressing attention is the Ryd On-Road Protocol (25, 26), which also adds addressing the blind spots to the right and left.

To address visual attention, which is the relevant attention in car driving (27, 28), the OT has to predict on what objects, or at least in which direction, the driver allocates his/her gaze. Commonly the only extra mirror on the windscreen in a driving assessment situation is set to assist the driving instructor to observe fellow road users. Consequently, eye contact cannot always be established between the OT assessor and the driver. Then, the only means of assessments left for the OT is to study the head movements of the driver. People without impairments tend to move their head rather than their eyes when they want to shift the focus of visual attention more than 5 × 5 degree (29), but otherwise keep their head still while viewing different objects within a visual field of <5 × 5 degrees. Consequently, with respect to objects that are within 5 × 5 degrees from previous fixation within the visual field, the assessing OT can, at best, only make an educated guess about the driver’s gaze direction, which risk jeopardising validity and reliability of the attention assessments. To complicate the matter even further, many of those assessed by OTs do in fact have some sort of physical impairment, in addition to cognitive and/or visual perceptual disabilities (30). For people with loco-motor impairments, e.g. PD, the inherent rigidity, akinesia and lack of motion following the disease (31) will further compromise the assessment by minimising their head movements. PD drivers have been found to have minimised neck and trunk rotation in observing traffic in T-junctions and roundabouts (7).

Whether visual attention is assessed by mere speculation or not, introducing possible reliability issues, has previously not been studied; most likely because it is hard to control for confounding factors in real traffic.
environments where these assessments take place (24, 27, 30, 32, 33). However, in a controlled environment, and with the actual eye moments concurrently established with the assessment in a mock-up situation, the accuracy of the OT assessment of attention, in addition to inter-rater reliability (IRR) could be addressed. Hence, in terms of correctly classified attention, the aim of the study was to establish the accuracy and the IRR of OT driving assessments of visual attention from the back seat. Furthermore, by establishing eye contact between the assessor and the driver through an additional mirror on the wind screen, the present study aimed to establish how much such an intervention would enhance the accuracy of the visual attention assessment.

Methods

Subjects
Two drivers with PD aged 56 and 59 with driving history: 37 and 40 yrs and six control drivers (Mean age: 49.8; driving history: 35 yrs on average) participated in the study. Hoehn & Yahr Stage of PD were 1.7 and 1.9 (34); and years of confirmed diagnosis were four and six. All participants recruited through convenience sampling wore corrective spectacles during the assessments. One OT-trained assessor (third author) did all the assessments and the IRR was tested with a third year medical student, not trained in driving assessments, but knowledgeable in the area of human body movements. Ethics approval to conduct the study was granted by the Human Research Ethics Committee of Curtin University (Approval number: OTSW-17-09). Upon arrival at the laboratory, informed consent to participate was obtained from all participants. Participants were also informed of the confidentiality of the study and their rights to withdraw their participation from the study at any time without any given reasons, with no consequences incurred.

Apparatus and procedures
The trials took place in the Curtin University Driving Rehabilitation Clinic in Perth, Western Australia. The subjects wore a head mounted eye tracker, Arrington ViewPoint Systems (35) shown in Fig. 3, while driving a fixed route in a PC-based STISIM fixed base driving simulator (36). The validity of the Curtin University STISIM driving simulator has been established through the assessment of driving performance of older adults (37). It was reported that there is a high transferability in the simulated and the on-road driving performance. The simulator consists of a mid-sized sedan (adjustable seat, brake and acceleration pedals and steering wheel) with an automatic transmission interface, as shown in Fig. 2. The experimental trials consisted of a continuous run of driving scenarios that included two-way and four-way roads metropolitan and country roads, intersections with and without stop signs and give-way signs. The drive took approximately 5–7 minutes to complete.

In addition, participants were required to respond to a secondary arithmetic task, previously used to study cognitive overloading in PD drivers (38, 39). It was presented on a screen displayed in front of the driver, as indicated by arrow 2 in Fig. 2. The drivers were asked to drive as they would normally do on the road, and at the same time attend to the secondary task. The task was to look at simple additions and press ‘Yes’ or ‘No’ on a knob on the steering wheel to indicate whether the suggested sum of the addition was correct or not. The reason for using this kind of visual secondary task was that, in order to respond correctly, the subjects had to foveate the numbers presented on the screen, since text and numbers cannot be read through peripheral vision (29). As shown in Figs. 1 and 2, the offset between the secondary task screen (arrow 2) and the focus of expansion in the central screen (arrow 4) was more than 5 × 5 degrees in the visual field.

For the fixation analysis, seven areas of interest (AOI) were defined, viz.: 1, Left side mirror; 2, Secondary task screen; 3, Interior rear view mirror; 4, Focus of expansion; 5, Right side mirror; 6, Speedometer and 7, Push buttons. As shown in this case, the arithmetic sum (indicated by arrow 2) was not correct and the subjects were supposedly pressing the ‘No’-button (indicated by arrow 7) as a response.

A centroid mode algorithm (13) fixation generation program built into the Arrington ViewPoint software (35) was set to recognise fixations, in which at least six consecutive data samples fell within a minimum of 1 × 1 degree of each other, providing a minimum fixation duration of 100 msec (14). Fig. 3 shows the Arrington ViewPoint eye tracker recording eye movements in 60Hz with a precision of 0.2 degrees.

Fig. 1. An example of a driving assessment situation. In the left seat the driving instructor equipped with dual commands an additional mirror. No extra mirror for the driving assessor sitting in the back seat. The three circles represent 1, 5, and 10 degrees of the visual field, based on the focus of expansion in this particular scene.
The head mounted eye tracker, presented in Fig. 3, was mounted and manually calibrated with a 16 point grid. Fixations were manually assigned to one of the seven AOI or as outside of them (missing data) for the following timeslots: 60 seconds to 61 sec. 70 to 71 sec. etc. until the pre-programmed drive was finished. A post-hoc manual video analysis, frame by frame, was made for the one second timeslots for each subject. All 245 fixation timeslots were possible to classify as within one of the seven AOI, and hence, no missing data were identified. However, several AOI could be registered for one timeslot, since a typical fixation duration is \( \sim 360 \) msec. (SD 220) with a skewness of 2.18 (27). In real traffic environments, fixation durations have been found to be typically shorter (14); on average 172 msec. (SD 62, skewness 2.49) in dense city traffic, and somewhat longer on rural roads, i.e. 196 msec. (SD 81, skewness 2.20). Consequently, over a time slot of one second, a driver can be expected to make 3–6 fixations on different objects in the visual field.

The time stamping was made on-line, since the actual time in seconds was presented in the lower centre part of the middle screen below (arrow 6 in Fig. 1) of the simulator, and thereby recorded by the eye tracker video. The visual attention of the drivers was manually assessed within a 1 second timeframe at intervals of every 10 seconds, starting at 60 seconds to 61 sec. 70 to 71 sec. etc. until the pre-programmed drive was finished. The assessor was given the chance to score multiple AOI during each one second time slot, but in reality it did not happen. This meant that 22–40 measurement points per subject were recorded on a scoring sheet by the assessor (third author). The assessor sat in a position resembling the position of an assessor sitting in the back seat of a car on the opposite side of the driver, i.e. the x, y, z-coordinates of the eyes were in the range of those that would appear if they were measured in a real on-road assessment.
For each subject, a spreadsheet was made where all fixations within the one second timeslots were allocated an AOI as defined by the eye tracking data, in addition to the OT assessor’s manual assessment of the visual attention of the driver in the same one second timeslot. Since several fixations, i.e., several AOI, were foveated during a one second timeslot, a perfect match of the OT manual AOI assessment to the AOI according to the eye tracking data was established when the OT manual assessment scored at least one of the fixed AOI during the same time slot.

The IRR was made while scoring AOI without a mirror when a PD driver drove. That particular drive comprised 33 data points, with additional 12 AOI points generated from the eye tracker data. The two assessors sat next to each other with no occlusion towards the simulator but were blinded from each other’s scoring. IRR was calculated as percentage of agreement.

Results
In total, 100 assessment points were identified on both the scoring sheets and with the eye tracker data in the condition without the mirror with the control drivers, plus yet another 33 assessment points with a PD driver. Another 72 assessments were recorded with a mirror creating eye contact between the assessor and the control driver, plus 40 assessment points with another PD driver. In total, 245 assessments were done using both the scoring sheet and the recorded eye movements, 133 without a mirror and 112 with the mirror. For the IRR, yet another 33 assessment points were added in the without mirror condition.

The total number of identified AOI by the eye tracker data was 366, i.e. 121 extra AOI were identified that did not match the manually scored AOI.

Without mirror, the overall accuracy was 56% when assessing the three control drivers [correctly classified ratio: 0.36, 0.96, 0.36], and for the three controls with mirror: 83% [correctly classified ratio: 0.88, 0.80, 0.82].

With respect to the PD drivers the numbers were substantially different. For the PD driver without mirror the accuracy was 94%, whereas for the PD driver with a mirror the accuracy was 90%.

With respect to the IRR, a 73% agreement was found, i.e. 24 out of 33, all but one of them in AOI 4; Focus of expansion. The last one was in AOI 2; the Secondary task screen. In none of the other five AOI, agreement in the remaining nine assessments points was found.

Discussion
The accuracy of the visual attention assessments from the back seat without the assistance of a mirror was low. On average, about half of every assessment point was accurate, a result that would not stand any serious scrutiny. However, with respect to the PD drivers, an almost total agreement was found without the mirror. This result could be viewed as a surprise, since we expect PD drivers to move their head less on the cost of moving their eyes. Further exploration of the data revealed, however, that in this particular case, only in three out of the 33 data assessment points timeslots, an AOI other than AOI 4 was present, i.e. Focus of expansion. Consequently, not only did the PD driver keep the head still, but also the movement of the eyes were minor. The corresponding ratio for PD driver number two, who drove with the mirror was five out 40 data assessment point AOI. PD is typically characterised by motor symptoms (31), cognitive deficits (40) in areas such as attention (41), memory (42), information processing (38) and executive functioning (43, 44) and by difficulties to engage in purposeful, self-directed, and self-serving behaviour (45). Mental inflexibility, slow reasoning and inability to self-monitor driving behaviour are hallmarks of executive functioning deficits (43, 44). Given these characteristics, it is easily anticipated that visual scanning patterns would be restricted. Consequently, the assessment task was very easy with respect to making an educated guess on which AOI they paid their attention to. Based on these pilot data, to add the mirror may not seem to be as important for this particular group of drivers. However, this needs to be further investigated on a large scale. OTs are increasingly more often involved in on-road driving assessments. However, these on-road assessment have been criticised for low validity and reliability (26), which is quite problematic given that a driver’s licence might be revoked based on the outcome of the test. To add an additional mirror is a way to improve the quality of visual attention part of the assessment. In real on-road assessments it is both cheap and physically feasible and it significantly raised the visual attention assessment accuracy, i.e. with almost a 50% increase in accuracy. Still, 83% accuracy is not acceptable when it comes to such important issues as the access to free and spontaneous mobility as offered by driving (46, 47). It is a well-known fact that driver cessation is related to an increase in depression and lack of participation in society (48). For many, cancellation of one’s driver license can lead to a major loss of control and independence (49). Unmotivated licence cancellation is thus a threat to public health from that perspective (50). To possibly restrict mobility by using a driving assessment protocol that takes visual attention into account and use it as part of the overall pass/fail assessment, without using at least a mirror is therefore, at best, malpractice. An 83% accuracy level indicates almost every fifth assessment is erroneous, which may have an impact on the safety aspect of the assessment, i.e. drivers who are unfit to drive continue driving because their problems with visual attention was not adequately assessed. The present study indicates
low or no validity of such an assessment procedure. In addition, it appears to have an insufficient IRR, further jeopardising the quality of a final pass/fail outcome.

If the final outcome of a driving assessment is dependent on the subcategory of a protocol assessing attention, we suggest the use of eye tracker based assessments to accurately determine the fixation points of the drivers. As mentioned, it can capture all the foveated objects in the visual field of the driver with a good precision, given that it is precisely calibrated. Since 3–6 different objects may be fixated per second, a registration rate that an assessor sitting in the back seat hardly can catch up with, the risk is that the crude manual OT assessments may miss fixations on crucial traffic objects that the driver did, regardless of how they moved their head. In addition, lots of information in traffic is processed through the peripheral vision (29, 30), as the driver is him/herself a moving object relative to other road users and roadside objects. An eye tracker can provide accurate information on fixations and visual search patterns, but eye movement data cannot rule out that a certain object has not been seen (as opposed to looked upon) by a driver (51, 52). Consequently, the OTs’ back seat observations on-road should not be a standalone visual attention assessment in driving assessments. Instead, we advocate a usage of eye trackers for further assessment in cases where there is doubt about a driver’s visual attention. However, with regard to feasibility, a minimum standard is a designated mirror for the assessor to make eye-contact with the driver.

**Limitations**

The participants who volunteered in the current study cannot be taken as representative of the older PD drivers population. The relatively small size of the monitor display of the driving simulator, together with the nature of the computer-generated stimuli from a stationary model car may limit the equipment to assess driving tasks that require complex visual perceptual abilities (37). This study was small with respect to the number of subjects, but 245 data assessment points is a substantial number of assessments, large enough to conclude that the accuracy of back seat visual attention assessments could be, and should be, questioned. However, further research is needed where the findings of this pilot study forms a basis.

**Conclusion**

If the final outcome of a driving assessment is dependent on the subcategory of a protocol assessing visual attention, we suggest the use of eye tracker based assessments to accurately determine the fixation points of the drivers. As a very minimum, an additional mirror should be used to establish eye contact between the assessor and the driver. Consequently, the clinicians’ observations on-road should not be a standalone assessment in driving assessments. Instead, we advocate a usage of eye trackers for further analyses and correlation in cases where there is doubt about a driver’s visual attention.

**Acknowledgements**

The study would not have been possible without the participants’ active involvement. The authors are indebted to the Sir Charles Gardiner Hospital and the Parkinson’s Association of WA for assisting the recruitment of participants and third-year medical student, Ola Redfors, for contributing to the IRR data collection. Thanks are also due to the editor and two anonymous reviewers for their constructive comments and suggestions.

**Conflict of interest and funding**

The authors have not received any funding or benefits from industry to conduct this study. This study was supported by a research grant from Curtin University, WA, Australia.

**References**


