

**MIND-WANDERING AND DRIVING: COMPARING THOUGHT REPORT AND
INDIVIDUAL DIFFERENCE MEASURES**

Heather E.K. Walker, Lana M. Trick
Department of Psychology, University of Guelph
Guelph, Ontario, Canada
Email: hrodd@uoguelph.ca

Summary: Mind-wandering is a cognitive state in which attention is diverted from the main task and towards more personal thoughts, which can interfere with performance. This study investigated differences in patterns of mind-wandering and driving performance measured during thought-probe versus post-task self-report conditions, and further differentiated based on individual differences in working memory—as measured by the *Operation Span (OSPAN)* and *Sustained Attention to Response Task (SART)*. Participants completed two 30-minute drives. Those in the thought-probe condition were asked whether they were thinking of driving; the proportion of trials where they answered “no” was used as the index of mind-wandering. In the post-task condition participants estimated the percentage of time they had mind-wandered during each drive. Speed, steering variability, headway distance, and hazard response time to a lead vehicle braking were also measured. Results showed that the magnitude of mind-wandering captured in the thought-probe condition was greater than in the post-task condition, though hazard response times were also faster despite greater mind-wandering reports. Higher *OSPAN* scores were associated with greater reports of mind-wandering, but only in the post-task condition. Conversely, in the post-task condition those with low *SART* scores responded slower to hazards than those with high scores; in the thought-probe condition these groups did not differ. Findings indicate a differential impact of report-type on participant experience, emphasizing the need for more covert measures of mind-wandering—e.g., eye-tracking or electroencephalography—that provide accurate estimates of task engagement but don’t interfere with task flow.

INTRODUCTION

Many drivers have had the experience of “coming to” in their driveway with little recollection of how they got from point A to B. This shift in attention away from a main task and towards one’s own thoughts is known as mind-wandering (Smallwood & Schooler, 2006), and there is already evidence that it can affect driving performance (e.g., He, Becic, Lee, & McCarley). Mind-wandering is often measured using in-task thought probes or post-task self-reports, but despite often being treated as equivalent it is unclear whether these techniques capture the same overall experience of mind-wandering or the associated changes in driving performance. If not, this could contribute to unexpected conflicts in the literature. In this study we examine this problem by carrying out an investigation of mind-wandering in a driving simulator, measuring the incidence of mind-wandering using two different methods: in-task thought-probes, and post-task reports. We also compare driving performance between these conditions to determine whether report type may secondarily affect driving behaviour. To further investigate if each report type is

measuring the same experience, we also compare patterns of mind-wandering and driving performance between those who score high versus low on two individual differences measures.

The study of mind-wandering and driving is not new. Research has revealed that drivers who are mind-wandering exhibit less complex eye movements and reduced horizontal visual scanning (He, Becic, Lee, & McCarley, 2011; Regan, Hallett, & Gordon, 2011), as well increased driving speeds and reduced headway distance (Yanko & Spalek, 2013; Walker & Trick, 2018)—all of which contribute to increased risk of collision. These effects can be further compounded by other factors such as familiarity with a route (Yanko & Spalek, 2013) or increased time spent driving (Walker & Trick, 2018)—both of which are a concern for commuters.

In many of these investigations, the most practical way in which to measure participants' mind-wandering has been to ask them through self-report. Two types of self-report commonly used throughout the mind-wandering literature include in-task thought probes and post-task self-reports (e.g., Smallwood & O'Connor, 2011). With probe-caught methods, auditory or visual cues prompt individuals to report whether they were mind-wandering in the past few moments, thus eliminating the need to always be aware of one's thoughts. However, probe-caught methods create an interruption—both in the flow of the task, and in the train of thought. Post-test surveys, on the other hand, are less intrusive but rely on individuals being able to identify and then later remember whether (and how often) mind-wandering occurred during the task. As such, there is a risk that mind-wandering will be underestimated because individuals may forget that mind-wandering occurred by the time the task is complete. Both approaches have advantages and disadvantages, and it is unclear whether the mind-wandering measured by these two techniques reflect identical experiences of off-task thought. If they do not—but are compared as if they were—are—this would cause confusion within the literature and could result in conflicting findings.

If probe-caught and self-report techniques are indeed capturing the same experience, then the magnitude of mind-wandering and pattern of changes in driving performance associated by individual differences in factors such as working memory (important in sustaining attention and coordinating action when attention is shared between two tasks, Watkins, 2008) should be similar between the two report-types. Two common measures of working memory are the *OSPAN* (*Operation Span*: Turner & Engle, 1989) and the *SART* (*Sustained Attention Response Task*: Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). The *OSPAN* measures ability to multitask and prioritize task goals over distracting stimuli (e.g., remembering an increasingly long list of random letters while at the same time carrying out simple mental calculations). The *SART* measures ability to maintain focused attention through a sequence of rapidly presented stimuli (e.g., numbers presented in rapid succession), responding and withholding response as instructed. It has been demonstrated that lower scores on these tasks are associated with greater instances of off-task thought, and it is hypothesized that those who perform poorly on these tasks are more likely to experience mind-wandering—as well as its associated performance deficits—in other day-to-day tasks as well (Kane & McVay, 2012).

In the present study we investigated mind-wandering and driving performance in a driving simulator under two self-report conditions: in-task thought probes, and post-task self-reports. Although they measure mind-wandering in different ways, results uncovered using these methods are often treated as equivalent in the literature, and so we wanted to examine the proportions of mind-wandering being reported by each approach. Because individuals may forget

that they were mind-wandering by the end of the task, we expected that lower magnitudes of mind-wandering would be captured by post-task self-reports. We also examined whether driving performance was also affected by how mind-wandering is measured. Interrupting the flow of the task to make a report versus carrying out an entire drive uninterrupted could have a differential impact on participants' driving performance, and so we suspected that driving performance would be better in the thought-probe condition as compared to the post-task condition. To further investigate whether each report type was measuring the same overall experience, we used the *OSPAN* and the *SART* to examine patterns of individual differences in mind-wandering and driving performance in each report type. Because both tasks measure overall working memory, it should be expected that patterns of mind-wandering and driving performance associated with high versus low scores on either of these tasks would be the same regardless of report type—if indeed they are measuring the same thing.

METHODS

Participants

Sixty-three licensed undergraduate students—a population more prone to mind-wandering (e.g., Zavagnin, Borella, & De Beni, 2014)—were recruited from the University of Guelph's psychology participant pool (38 female, M age = 19.19, SD = 2.13), and received course credit for their participation. All had at least a G2 license or better, which in Ontario's graduate licensing program signifies that the driver is entitled to drive unsupervised. 70% of all participants had a G2 license, with the remaining possessing a full (G) license.

Apparatus and Stimuli

This study was carried out using a fixed-base *Oktal* driving simulator, consisting of a *Pontiac G6* convertible surrounded by 300° of viewing screens. The vehicle is equipped with the standard controls, including an automatic transmission. To further contribute to a realistic driving experience, the vehicle is also equipped with speakers and vibration transducers to simulate the sounds and vibrations experienced during acceleration, as well as force feedback in the steering wheel to simulate feedback experienced when steering a vehicle.

Two similar highway drives were created, which took 30 minutes to complete at 100 km/h. Participants followed behind a lead vehicle throughout each drive and were told to maintain an appropriate following distance. Near the end of each drive the lead vehicle braked unexpectedly, and participants had to brake to avoid a collision. Hazard response time (RT) was calculated both as the time (in ms) between the lead vehicle braking and when participants lifted their foot off the gas pedal (gas RT), and the time between the lead vehicle braking and when participants applied brake pressure (brake RT). Other measures of driving performance included driving speed, standard deviation of lateral position (SDLP: standard deviation of the distance between the center of the vehicle and the center of the lane), and headway distance. These were measured during straight and uneventful highway portions where mind-wandering would likely occur.

Two methods were used to measure mind-wandering. In the thought-probe condition, at six points throughout each drive a pre-recorded voice was played asking, "Are you thinking about driving?" On the dashboard above the center console were two buttons labelled "Y" and "N". If participants were thinking of driving, they pressed "Y". If not, they pressed "N". Of interest was

the proportion (%) of total probes to which participants responded “N” (i.e., were mind-wandering). In the post-task condition, participants were asked to estimate the proportion (%) of time they felt they had spent mind-wandering during each drive.

Individual differences were assessed using two measures of executive working memory. In the *OSPAN*, participants were presented with math equations and first had to determine whether the given answer was true or false, after which they were shown a letter which they were told to hold in memory. After several math equation and letter pairs, they were asked to recall the letters in the order they were presented. The number of equations and thus letters to be recalled were increased until participants could no longer recall the letters correctly on 3 successive trials. Scoring was based on the total number of correct answers. For the *SART*—a go/no-go task—participants’ task was to tap an *iPod* screen every time a number appeared (“go” items), *except* when that number was a “5” (“no-go” item). Of interest was total test accuracy, a percent-correct score which factors in correct responses to both “go” (respond) and “no-go” (withhold) items.

Procedure

Participants were randomly assigned to either the thought-probe or post-task condition. Upon arrival for the experiment, participants were presented with the informed consent information. To avoid making them wise to the full purpose of the experiment, participants were not informed of what to expect during the drives (i.e., that the lead vehicle would be braking), apart from the fact that they would be responding to yes/no questions during the drives (thought-probes) or rating their experience after the drives (post-task reports). However, they were not informed as to the specific content of these questions. After filling out intake questionnaires that collected basic demographics and driving experience information, participants completed a five-minute practice drive to become familiar with the feeling of being in the simulator and with piloting the vehicle. They were then administered both the *OSPAN* and the *SART*. Following this they completed the two experimental drives, the order of which was counterbalanced across participants.

RESULTS

Design. This study involved two between-subject’s factors: report type (thought-probe, post-task), and executive working memory (high, low), the latter of which was determined using a median split of both *OSPAN* and *SART* scores. Data analyses involved analyses of variance (ANOVA, with Type III Sums of Squares used when factorial analyses were conducted). Effect sizes were measured using eta-squared (η^2) and partial eta-squared (η_p^2), which represent the proportion of variance accounted for by each experimental factor. When tests of means were performed, a *Bonferroni* correction was used to correct for multiple comparisons.

Report type. One-way ANOVA were used to compare mind-wandering and driving performance between the two report types. The proportion of mind-wandering reported in the thought-probe condition was higher than estimates from the post-task condition ($F(1,61) = 4.61, p = .036, \eta^2 = .07, M_{diff} = 14.71\%$) (Fig. 1 left). Comparisons of driving performance revealed that gas RTs were higher in the post-task condition than in the thought-probe condition ($F(1,51) = 4.66, p = .036, \eta^2 = .084, M_{diff} = 389.5\text{ms}$) (Fig. 1 middle). A similar pattern emerged for braking RTs, though it was only marginal ($F(1,49) = 3.51, p = .067, \eta^2 = .067, M_{diff} = 387.15\text{ms}$) (Fig. 1 right). Other comparisons of driving performance revealed no difference.

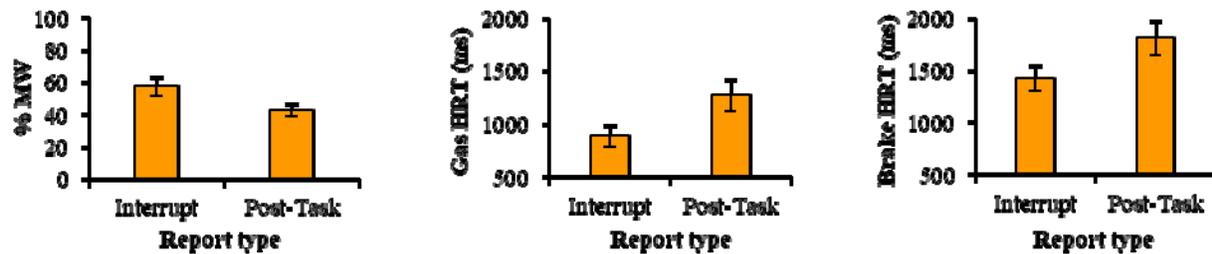


Figure 1. Reported mind-wandering (Left), gas RT (Middle), and brake RT (Right) for each report condition. Standard error bars included

Individual differences. To confirm that differences between report conditions were due to experimental manipulation and not to pre-existing differences between groups, comparisons of overall *OSPAN* and *SART* performance between each group were carried out. Results showed no difference in participants' overall *OSPAN* scores ($F(1,62) = 2.39, p = .128, \eta^2 = .038$) or *SART* scores ($F(1,61) = 1.02, p = .318, \eta^2 = .017$). Participants scores were then divided into high and low *OSPAN* or *SART* performance groups based on a median split of overall scores for each test (*OSPAN* median = 18 (out of 25); *SART* median = 95.2% accuracy).

In contrast to overall comparisons between the report conditions, when the impact of individual differences in working memory on reports of mind-wandering were examined, separate one-way ANOVA were used. In the post-task condition, individuals with high *OSPAN* scores estimated mind-wandering marginally more often than those with low scores ($F(1,29) = 3.86, p = .059, \eta^2 = .12, Mdiff = 14.47\%$) (Figure 2). The proportion of mind-wandering reported in the thought-probe condition followed a similar pattern, though this difference was not close to significance ($F(1,30) = 1.48, p = .23, \eta^2 = .047, Mdiff = 14.29\%$) (Fig. 2 left). When comparisons were made using high and low *SART* scores, no differences were found.

2 X 2 factorial ANOVA were used to examine differences in driving performance, with report type (thought-probe, post-task) and *SART* or *OSPAN* score (high, low) as between-subjects' factors. We found an interaction between report type and *SART* scores on braking RTs ($F(1,45) = 4.70, p = .014, \eta_p^2 = .17$) (Fig. 2 middle). In the post-task condition those with low *SART* scores had marginally higher braking RTs than those with higher scores ($Mdiff = 633.17\text{ms}, SE = 323.22\text{ms}, p = .056$), while in the thought-probe condition there was no difference between the two ($Mdiff = 80.0\text{ms}, SE = 227.06\text{ms}, p = 1$). Furthermore, low scorers in the post-task condition had higher braking RTs than low scorers in the thought-probe condition ($Mdiff = 770.60\text{ms}, SE = 269.74\text{ms}, p = .008$), whereas there was no difference between the two conditions for those with higher scores ($Mdiff = 57.50\text{ms}, SE = 275.02\text{ms}, p = 1$). An interaction between report type and *SART* scores was also found for gas RTs ($F(1,47) = 3.62, p = .034, \eta_p^2 = .134$) (Fig. 2 right). Specifically, in the post-task condition those with low *SART* scores had higher gas RTs than individuals with higher scores ($Mdiff = 645.06\text{ms}, SE = 294.94\text{ms}, p = .023$), while in the thought-probe condition there was no such difference ($Mdiff = 91.0\text{ms}, SE = 184.10\text{ms}, p = 1$). Additionally, low scorers in the post-task condition had higher gas RTs than low scorers in the thought-probe condition ($Mdiff = 674.06\text{ms}, SE = 259.33\text{ms}, p = .007$), while those with high scores did not differ between the two report conditions ($Mdiff = 120.0\text{ms}, SE = 144.25\text{ms}, p = 1$). Other comparisons of driving performance revealed no differences.

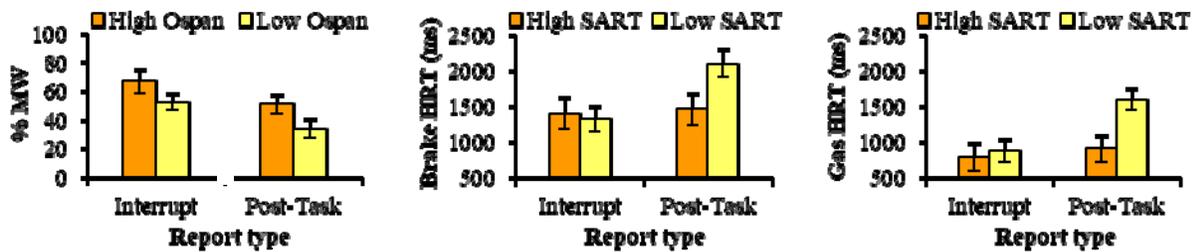


Figure 2. Individual differences in mind-wandering as indicated by OSPAN performance (Left), individual differences in brake RT (Middle) and gas RT (Right) as indicated by SART performance, for each report condition. Standard error bars included

DISCUSSION AND CONCLUSIONS

As predicted, the magnitude of mind-wandering captured in the thought-probe condition was greater than that estimated by post-task self-reports. This is not to say that less mind-wandering was experienced by those in the post-task condition, only that individuals may not have been aware of it or simply forgot, and so a portion of episodes went unreported. Looking to measures of driving performance, only hazard RTs differed between the two report conditions. Hazard RTs were faster in the thought-probe condition, even though more mind-wandering was reported. Two explanations are possible: interrupting thought-probes might have reduced the impact that extended mind-wandering had on driving behaviour, and so performance was better. Alternatively, if individuals in the post-task condition are mind-wandering more than is being reported, it would contribute to the longer RTs that were observed.

For comparisons involving individual differences, only the *OSPAN* was associated with differences in reported mind-wandering. Those with high scores reported more-mind-wandering than those with low scores, though only to a marginal degree in the post-task condition. It is certainly possible that these individuals mind-wandered more than those with low scores. However, because the *OSPAN* measures ability to keep items in memory while completing a second task, it is also likely that they were simply better at recognizing they were mind-wandering. Conversely, only the *SART* was associated with differences in driving performance. Those with low scores experienced longer hazard RTs, though only in the post-task condition. Again, it is possible that thought-probes brought attention back on-task just enough to eliminate noticeable differences between the two groups. Overall, patterns of mind-wandering and driving performance associated with high and low scores on each task were not the same between the two report conditions, as would be expected if each report type were surveying the same experience. Instead, it suggests the opposite, that the behaviours observed when using thought-probes are not necessarily the same as those observed when using post-task reports.

Although no firm conclusions should be made about the strengths of each report type, it could be that thought-probes are ideal when accurate estimates of mind-wandering are required, whereas post-task reports are ideal when performance is the focus. Nevertheless, these results imply that conclusions made in experiments using one measurement type are not necessarily generalizable to other experiments or contexts—though they are often treated as such in the literature—and so caution and scrutiny should be used when interpreting such findings. There are other general implications as well. Care should be taken when choosing a measurement approach, as certain methods may not accurately represent how much mind-wandering has occurred. Furthermore,

measures should be chosen carefully to minimize any secondary impact on associated behaviours that are also being measured. This highlights the need for more covert measures of mind-wandering—for example, eye-tracking or electroencephalography (EEG)—that provide accurate estimates of attention towards the task but don't interfere with task flow.

ACKNOWLEDGEMENTS

We would like to thank the *Natural Sciences and Engineering Research Council of Canada* for funding this investigation, as well as the *Canadian Foundation for Innovation* and the *Ontario Research Fund*, who co-supported the upgrades to the University of Guelph's simulator facility.

REFERENCES

- Barragan, D., Roberts, D. M., Chong, S. S., & Baldwin, C. L. (2017). Comparing methods of detecting mind wandering while driving. *9th international driving symposium on human factors in driver assessment, training and vehicle design* (pp. 79–86).
- He, J., Becic, E., Lee, Y.C., & McCarley, J. S. (2011). Mind wandering behind the wheel: Performance and oculomotor correlates. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *53*(1), 13–21.
- Kane, M. J., & McVay, J. C. (2012). What mind wandering reveals about executive-control abilities and failures. *Current Directions in Psychological Science*, *21*(5), 348–354.
- Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis and Prevention*, *43*(5), 1771–1781.
- Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). 'Oops!': Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747–758.
- Smallwood, J., & O'Connor, R. C. (2011). Imprisoned by the past: unhappy moods lead to a retrospective bias to mind wandering. *Cognition & emotion*, *25*(8), 1481–1490.
- Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of memory and language*, *28*(2), 127–154.
- Walker, H. E., & Trick, L. M. (2018). Mind-wandering while driving: The impact of fatigue, task length, and sustained attention abilities. *Transportation Research Part F: Traffic Psychology and Behaviour*, *59*, 81–97.
- Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, *134*(2), 163–206.
- Yanko, M. R., & Spalek, T. M. (2013). Driving with the wandering mind: The effect that mind-wandering has on driving performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(2), 260–269.
- Zavagnin, M., Borella, E., & De Beni, R. (2014). When the mind wanders: Age-related differences between young and older adults. *Acta Psychologica*, *145*, 54 – 64.