Influence of automation on mind wandering frequency in sustained attention

Jonas Gouraud, Arnaud Delorme, Bruno Berberian

To cite this version:

Jonas Gouraud, Arnaud Delorme, Bruno Berberian. Influence of automation on mind wandering frequency in sustained attention. Consciousness and Cognition, 2018, 66, pp.54-64. 10.1016/j.concog.2018.09.012 . hal-02314045

HAL Id: hal-02314045
https://hal.science/hal-02314045

Submitted on 11 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Influence of Automation over Mind Wandering Frequency in Sustained Attention

Jonas Gouraud[1], Arnaud Delorme[2,3], Bruno Berberian[1]

1. System and Information Processing Department, ONERA, Salon de Provence, France

2. Center of Research on Brain and Cognition (UMR 5549), CNRS, Toulouse, France

3. Swartz Center for Computational Neurosciences, University of California, San Diego, California, USA

Correspondence concerning this article should be addressed to Jonas Gouraud, ONERA Salon de Provence, 13661, Salon Air, France. Contact: contact@jonasgouraud.com

Abstract

Recent evidences showing that mind wandering might fill the time saved by automation are particularly worrying when taking into account the negative effect of mind wandering on short-term performance. 17 participants performed an obstacle avoidance task under manual and automated conditions in 2 sessions lasting 45 minutes each. We recorded attentional probes, oculometry and answers to the Task Load Index after each session. Subjects perceived the manual condition as more demanding than the automated one. We highlighted a significant influence of automation on the mind wandering frequency after some time. Multiple phenomena may play a role, such as complacency and decoupling from the task at hand. Pupil diameter decreased during mind wandering versus focus periods, with a stable amplitude. Mind wandering knowledge could be used in a near future to characterize and quantify an operator's state of mind regarding automation related problems.

Keywords: mind wandering; automation; vigilance; oculometry; complacency; decoupling

Introduction

In order to continuously improve system safety, the critical systems industry makes extensive use of automation (Baxter, Rooksby, Wang, & Khajeh-Hosseini, 2012; Parasuraman, 1987). In cockpits (Wise, Tilden, Abbott, Dyck, & Guide, 1994), in cars (Naujoks, Purucker, & Neukum, 2016), and in power plant consoles (Cummings, Sasangohar, Thornburg, Xing, & D’Agostino, 2010), automation has been introduced to increase performance and respond to new safety requirements. Unfortunately, while implementing higher levels of automation indeed improves the efficiency and capacity of a system, it also creates new challenges for human operators. Particularly, the externally imposed task to maintain sustained attention focused for long periods of time in low probability environments causes progressive vigilance decrement – or invigilance increment (Hancock, 2013) – preventing efficient automation supervising (Amalberti, 1999). As targets are hidden – naturally, voluntarily or because of poor display design –, the task to detect and react to these targets is often stressful and increasingly difficult (Mackworth, 1948). These problems result in out-of-the-loop (OOTL) performance problem, referring to a performance decrease whenever attempts are made to regain manual control after a critical system failure.

Such problems have been studied in laboratories (Endsley & Kiris, 1995; Sarter, Woods, & Billings, 1997), but are also regularly reported in operational conditions. Mosier and collaborators (1994) examined NASA’s Aviation Safety Reporting System (ASRS) database and reported that 77% of the incidents involved an over-reliance on automation leading to a probable vigilance failure. Similarly, Gerbert and Kemmler (1986) studied German aviators’ anonymous responses to questionnaires about automation-related incidents and pointed out failures of vigilance as the largest contributor to human error. Several studies showed that efficient sustained attention for hours cannot be achieved (Cabon, Coblentz, Mollard, & Fouillot, 1993; Mackworth, 1948; Methot & Huitema, 1998).

Such a context may favor the occurrence of mind wandering (MW) episodes. MW is a family of experiences relating to the mind’s tendency to engage in thoughts unrelated to the here and now (Smallwood & Schooler, 2006). It is an ubiquitous phenomenon that can be intentional or spontaneous (Golchert et al., 2016; Seli, Risko, & Smilek, 2016), be guided or unguided (Smallwood, 2013), emerge when performing a task or at rest (Smallwood, Baracaia, Lowe, & Obonsawin, 2003) while its ignition point can be triggered by the environment or generated internally (McMillan, Kaufman, & Singer, 2013; Smallwood & Schooler, 2006). In the following paper, we focused on MW when performing a task without discriminating other dimensions. MW is more likely to occur in monotonous environments (Eastwood, Frischen, Fenske, & Smilek, 2012), or when operators perform familiar (Bastian et al., 2017) or long tasks (Smallwood & Schooler, 2015). Its occurrence favors a decoupling from the ongoing task at perceptual and stimuli processing levels (Kam et al., 2012; Schooler et al., 2011), which can be seen both on behavioral and physiological data. Reading tasks were particularly used to uncover the influence of MW over oculometric markers like blink frequency (Smilek, Carriere, & Cheyne, 2010b), fixation duration and saccade frequency (Uzzaman & Joordens, 2011). In simulators, Yanko and Spalek (2014) studied MW influence over driving performance. They observed a longer reaction time to unexpected events, a shorter headway distance and a higher velocity. Their results were corroborated by other studies in driving environments (Dündar, 2015; He, Becic, Lee, & McCarley, 2011; Lerner, Baldwin, Higgins, Lee, & Schooler, 2015).

Given that MW diverts an operator’s attention from his primary task, it could play an important role in vigilance failures observed in highly reliable automated environments. Casner and Schooler (2015) studied the impact of automation on MW in an aeronautical context. Their results on 16-minute sessions did not show a significant correlation between automation and the frequency of MW reports. However, the propensity to mind wander appeared to increase when everything seemed under control. Supervising ultra-reliable systems could encourage operators to decrease cognitive resources allocated to the monitoring task. In that context, time saved by automation, which should normally be used for other productive tasks and for monitoring, could instead be filled by task-unrelated thoughts. Operators in such a state would not be prepared to regain manual control over the system in response to rare critical events. Such analysis is already considered in the debate regarding the origin of the vigilance decrement (Fraulini, Hancock, Neigel, Claypoole, & Szalma, 2017; Pattyn, Neyt, Henderickx, & Soetens, 2008; Thomson, Besner, & Smilek, 2016), recent evidences showing that both phenomena share many features (Gouraud, Delorme, & Berberian, 2017).

We believe automation might influence MW during longer sessions within ecological environments. We think that this impact may be observable on the MW frequency, as well as on the physiological markers of MW. Our experiment addresses these hypotheses.

Material and methods

Participants

17 participants (5 female) performed the experiment (age ranging from 21 to 42 years old; *M* = 27.3, *SD* = 6.0). The participants enrolled in this study were volunteers from our company (ONERA organization). All participants had normal or corrected-to-normal visual acuity. All participants signed a written declaration of informed consent. The protocol was conducted in accordance with the Declaration of Helsinki.

Task

Environment. We used the LIPS (Laboratoire d’Interactions Pilote-Système, or Pilot-System Interactions Laboratory) environment developed at the ONERA organization to program our experiment (see Figure 1). An unmanned air vehicle (UAV) depicted as a plane seen from above stayed at the center of a 2D radar 22-inch screen and moved following waypoints arranged in a semi-straight line with clusters of obstacles along the way (every 45s on average). Each cluster contained between 1 to 5 obstacles, including one on the trajectory. The participants were instructed to control the movements of the UAV to avoid obstacles. The LIPS environment includes a physics engine to reproduce convincing *Rafale* aircraft motion behavior. The LIPS was displayed on a screen within the SIMPIT environment shown in Figure 1.

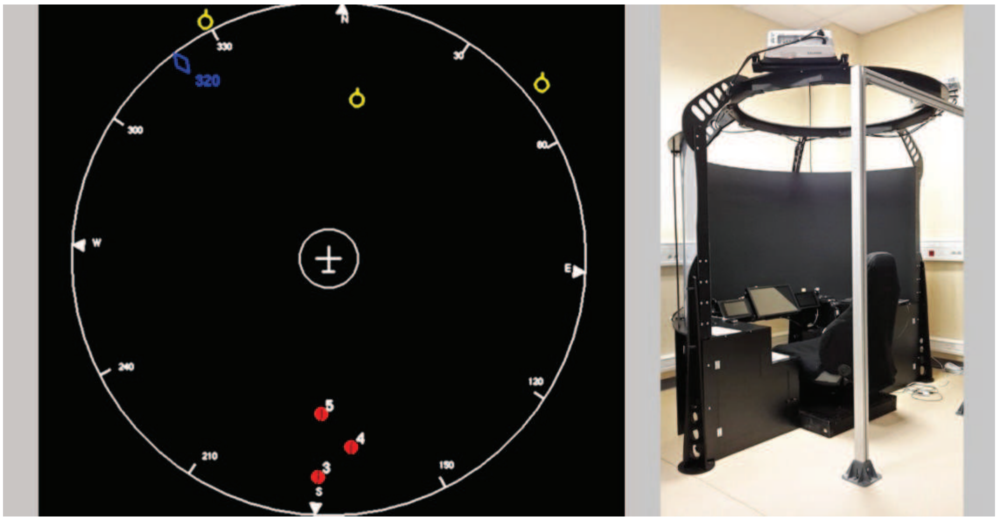


Figure 1. Screenshot of the LIPS interface and the environment. One of the screen is used for the task and the other one for questionnaire probes. For the task, the plane in the center is static and the surround (yellow and red numbered symbols) are moving. During, left and right avoidance maneuvers, again the plane is static and the background is rotated.

MW probes. Python 3.6 was used to program mental probes. On average every 2 minutes, the probe appeared on a secondary 10-inch screen next to the main screen. For technical reasons, the obstacle avoidance task was not paused when the probe was displayed. Participants were asked to fill it as soon as it appeared, and any successful or failed trial during this interval would not be taken into account. Participants were informed that the probe was not part of the evaluation to lower the impact of instructions over their natural propensity to mind wander. Participants were required to answer the following questions (originally in French, see Figure 2): “When this probe appeared, where was your attention directed?” Answers could be “On the task” (e.g., thinking about the next obstacle, the decision to make, the incoming waypoint), “Something related to the task” (e.g., thinking about performance, interface items, last trial), “Something unrelated to the task” (e.g., thinking about a memory, their last meal, or a body sensation, hereafter defined as MW) or “External distraction” (e.g., conversation, noise). The preceding examples were given to participants to illustrate each category. We were primarily interested in reports of being “On the task” and MW reports. Reports of thoughts “Related to the task” were integrated to avoid participants to report MW when thinking about their performance (Head & Helton, 2016). Noises were integrated to avoid participants to report MW if they were focused on any external signal.

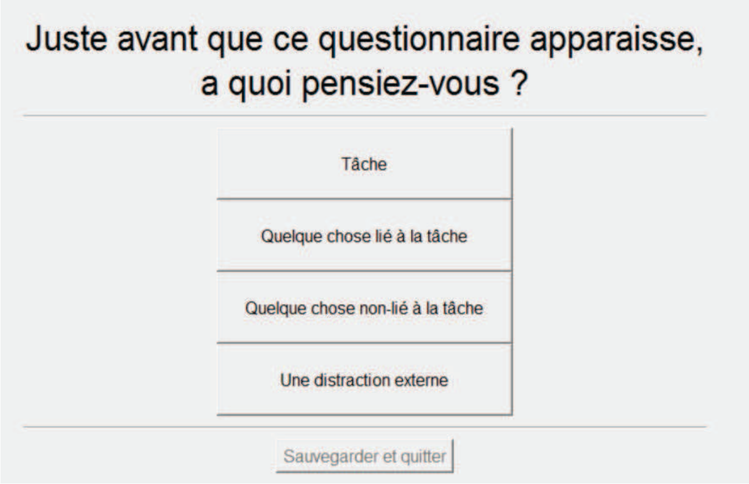

Juste avant que ce questionnaire apparaisse, a quoi pensiez-vous ?

Figure 2. Screenshot of the French MW probes. The question is “When this probe appeared, where was your attention directed?” Answers could be “On the task”, “Something related to the task”, “Something unrelated to the task” or “External distraction”

Conditions. Two conditions were proposed. The first one was the “manual” condition and required participants to manually avoid obstacles. The system detected obstacles on the trajectory 13s before impact. Then, an orange circle appeared around the UAV and the participant could initiate an avoidance maneuver. Participants were able to choose the way in which they wished to avoid the obstacle by clicking on “*Evitement Gauche*” (left maneuver) or “*Evitement Droite*” (right maneuver). Once they clicked, the simulator turned the trajectory of the UAV on the chosen side, following a predefined angle. Each obstacle had a safe circle similar to that of the UAV (see Figure 1). A collision warning – i.e., an orange circle around both the UAV and the obstacle with a message “Collision” – was displayed if the UAV safe circle penetrated inside the obstacle safe circle. A trial with a collision warning triggered was marked as failed. To resume the initial trajectory, they had to click on the “*Retour trajectoire*” (return to original trajectory) button. If no action was taken within 16 seconds after the first change in trajectory, the aircraft automatically resumed the trajectory and the trial was marked as failed.

The second condition was the “automated” condition. Participants were required to monitor the system avoiding obstacles. They had to click on an “*Acquittement*” (acknowledgement) button to acknowledge automated avoidance decisions as soon as they saw it – twice per trial, once to acknowledge avoidance of the object and once to acknowledge the return to normal trajectory after avoiding the object. A feedback message was displayed to the participants. The acknowledgement ensured that participants would have the same motor input under both the manual and the automated conditions. If participants detected an automation error, i.e. choosing the wrong side for avoidance trajectory, they were instructed to click on the button “*Changement d’altitude*” (change altitude) so that the UAV would perform an emergency descent. A feedback message was displayed in that case as well. The altitude change ensured that participants were facing a supervision task.

Procedure. Participants were explicitly instructed that detection accuracy was more important than speed in button clicks. Each participant performed the two conditions on two separate days in a counterbalanced way. Each day started with an explanation of the task, followed by a 10-minute training period and a 45-minutes session under the proper condition. Each session contained 60 clusters of obstacles. Each cluster was considered a trial. They were separated by 45 seconds on average. 20 probes were answered under each condition. The distribution of probes was not correlated with events on the obstacle avoidance task in order to avoid performance to influence MW reports (Head & Helton, 2016). The automated condition included 8 conflicts with a probe within the 10-seconds interval following the conflict. Participants encountered one system error (where they had to click on the "*Changement d'altitude*" button) during training, and another during the automated condition at the end of the third block. Under the manual condition, participants encountered at the end of the third block a conflict impossible to avoid. Both the automation error and this conflict were not followed by an attentional probe for at least 10 seconds after.

Data recording

MW Probes. Comma Separated Value (CSV) text files were used to store all answers. The exact appearance time was saved along with each answer, in order to synchronize probes data with the pupillometric signal.

Post-task questionnaire. We used a validated French version of the NASA Task Load Index (TLX) questionnaire to evaluate the required amount of cognitive resources – equated as workload – along several dimensions (Cegarra & Morgado, 2009; Hart & Staveland, 1988). This questionnaire includes questions pertaining to mental load, time pressure, physical strain, effort, frustration, and perceived performance. Participants were asked to answer each question using a horizontal line, ranging from 0 to 20. Although a

TLX questionnaire completed at each block would allow precise workload monitoring, we believe that MW would have been artificially lower due to the disruption. Therefore, the TLX was only filled at the end of each session.

Oculometry. Oculometric data was recorded using the hardware SmartEye Pro 3.0 and the software SmartEye 6.2.4. The system included 2 infrared illuminators and 3 cameras (120Hz) placed above the screen to avoid any direct contact with the participant (see Figure 1). Gaze calibration was performed using a 4-point grid.

Performance. We recorded button clicks throughout both conditions. Each button click was saved along with its timestamp within a CSV text format by the LIPS environment.

Data Analysis

MW probes. Participants' clicks and probe answers were saved in CSV text format. We used R-Studio and R 3.4.1 (R Core Team, 2016; RStudio Team, 2015) to analyze the data.

Oculometry. The 10 seconds preceding each probe were extracted from oculometric data. This period length is in line with the literature investigating MW and oculometric markers (Bixler & D'Mello, 2014, 2015; Franklin, Broadway, Mrazek, Smallwood, & Schooler, 2013; He et al., 2011). Extracts before "On the task" and "Something related to the task" were classified as "Focus" to avoid any influence of poor performance on subsequent attentional reports (Head & Helton, 2016). Extracts before "Something unrelated to the task" were classified as "MW". Extracts before "External distraction" were discarded as noise.

We performed filtering on pupillometry using the R package reshape (Wickham, 2007), psych (Revelle, 2017), ggplot2 (Wickham, 2009, p. 2) and robfilter (Fried, Schettlinger, & Borowski, 2014). Filtering was done in two passes, following a method already used in the literature (Grandchamp, Braboszcz, & Delorme, 2014). Firstly, we filtered the signal. Pupil diameter had to be between 1 and 10 mm (due to the physical limits of pupil diameter, see (Lemercier, 2014), had to be less than 80% different from the preceding value (due to pupil dynamic limits) and had to be of a quality (computed by the SmartEye software) over 0.01. Extracts were discarded if their resulting pupil diameter series consisted of less than 70% compliant values. The proportion of extracts excluded due to low quality (9.6%) is in line with that excluded in other investigations (Smallwood et al., 2011). Resulting extracts were completed using basic linear interpolation. A second filtering pass was applied with a median filter (moving window of 50 frames). Finally, the data of each participant were normalized by subtracting the mean and dividing by the root mean square of all good-enough quality extracts for this participant.

Fixations, saccades and blinks were computed by the SmartEye Pro software. Blinks were computed using sliding windows of 700ms. Saccades were defined in SmartEye Pro parameters as gaze velocity over 35 deg/s. Saccades were limited to 200ms. Fixations were frames where the gaze velocity remained below 15 deg/s.

Performance. Performance was assessed by determining if participants clicked when they were required to do so. Reaction time were computed by comparing participants button click time delay in the manual condition to the moment the system detected an obstacle, and in the automated condition to the time at which they acknowledged each automation decision.

Results

Mind Wandering Frequency

We split the 45-minute sessions into 4 blocks lasting 10-minutes and containing 5 reports each. MW propensity was calculated as a percentage of all reports in the block (see Figure 3). Participants reported MW episodes for almost half of the probes (M = 49%, SD = 30%). This rate is consistent with previous studies on the subject (Kam et al., 2011; Smallwood & Schooler, 2006, 2015). Each participant reported on average 4% "external distraction" thoughts in each session (SD = 5). Such a low rate justified discarding "external distraction" reports as noise without thwarting subsequent analysis.

We used the ezANOVA function (Lawrence, 2016) to perform a two-way repeated measure ANOVA. We entered time (block) and level of automation (condition) as independent variables. We used the MW frequency as a dependent variable. Mauchly's test indicated that the assumption of sphericity had been verified for the main effect of block, $W = 0.64$, $p = .251$, and block × condition, $W = 0.90$, $p = .906$. There were significant main effects of time over the MW frequency, $F(3, 48) = 8.88$, $p < .001$, as well as of the level of automation over the MW frequency, $F(1, 16) = 12.67$, $p = .003$. There was also a significant interaction effect between the time and the level of automation, $F(3, 48) = 5.22$, $p = .003$. Without specific *a priori* predictions on the evolution of MW frequency through time, we conducted Tukey's post-hoc tests on the model including Block variable for each condition separately. We used the *glht* (Hothorn et al., 2017) and *mes* (AC Del Re, 2014) functions. For the manual condition, all differences were non-significant (p > .366). For the automated condition, the third and fourth blocks had significantly higher MW frequency compared to the first block, p = .001, d = 0.54 and p = .003, d = 0.32, respectively. Similarly, the blocks 3 and 4 had significantly higher MW frequency compared to the block 2, p = .007, d = 0.12 and p = .016, d = 0.12, respectively.

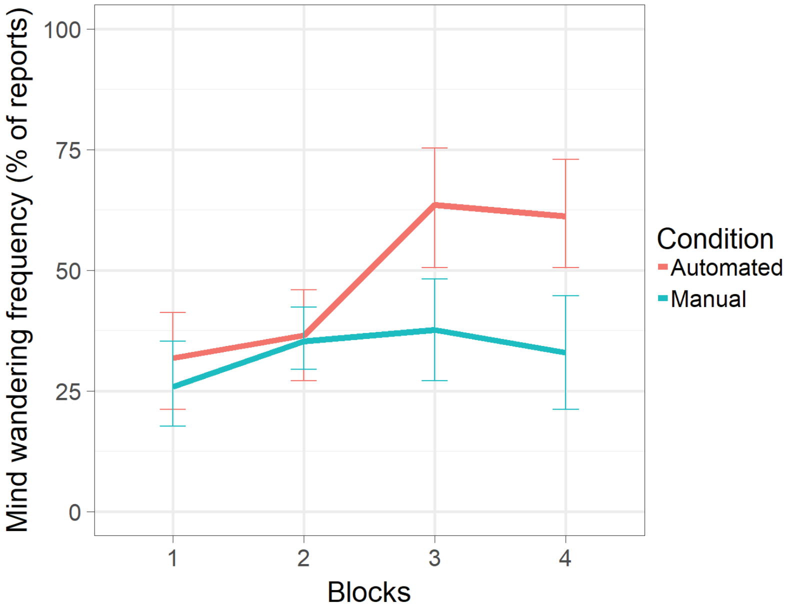


Figure 3: MW frequency evolution according to the condition. Error bars show the 95% confidence intervals based on bootstrap

NASA TLX Scores

Each participant filled 2 TLX questionnaires (one after each session). The mean score for each TLX of each subject (see Figure 4) varied substantially (ranging from 2 to 14.17, M = 5.81, SD = 2.44). Shapiro-Wilk's test indicated that the assumption of normality had been violated for the TLX values, W = .921, p = .012. Therefore, we used

Wilcox's robust version of the *t*-test proposed in the WRS2 package (Mair, Schoenbrodt, & Wilcox, 2017). On average, participants perceived that the automated (*M* = 4.93, *SD* = .50) condition required more cognitive resources than the manual (*M* = 6.68, *SD* = .61) condition, *t*(10) = -3.35, *p* = .007, *d* = 0.78. TLX scores show that our automated condition succeeded in lowering workload.

Figure 4: Normalized pupil diameter. Evolution during the 30-second interval preceding probes display – the grey part of the signal is used for computation

Oculometry

Influence of MW over oculometric measures. Oculometric measures were first analyzed using the 10 seconds preceding each probe. We used the lmer function (Bates, Maechler, Bolker, & Walker, 2015, p. 4) to perform a linear mixed-effect analysis despite missing values – "external distraction" reports and bad quality extracts excluded. As random effects, we had intercepts for subjects. Visual inspection of residual plots did not reveal any obvious deviations from normality or homoscedasticity. P-values were obtained by likelihood ratio tests using ANOVA on the full model against the models with no fixed effect, with only block, with block and condition, and with block and condition and their interaction. The results are shown in Table 1. On average, participants showed a significantly smaller pupil during MW episodes (see Figure 4). There was no main effect of MW on other markers.

Table 1: Comparison of oculometric measures during MW and focus episodes

Parameter	**MW values**		**Focus values**		**Mental State model**	
	M	*SD*	*M*	*SD*	$\chi^2(1)$	p-value
Pupil size (mm)	4.90	0.97	5.05	0.97	**7259**	**<.001**
Saccade frequency (sacc/s)	3.92	2.36	3.89	2.39	0.07	.795
Mean fixation duration (s)	$2.87*10^{-1}$	$6.65*10^{-1}$	$3.22*10^{-1}$	$6.52*10^{-1}$	0.08	.774
Blink frequency (blink/s)	$6.90*10^{-2}$	$1.10*10^{-1}$	$5.43*10^{-2}$	$9.81*10^{-2}$	2.09	.148

Influence of time and automation on oculometric differences. We looked for any influence of time or automation over the pupillometric differences previously observed between MW and focus periods. We used the lmer function to perform the linear mixed-effect analysis, as in the previous paragraph. As fixed effects, we entered time (block), level of automation (condition) and their interaction. As random effects, we had intercepts for subjects, as well as by-subject random slopes for the effect of block and condition. Visual inspection of residual plots did not reveal any obvious deviations from normality and homoscedasticity. Results are gathered in Table 2. Pupillometric difference remained stable through time and condition.

Table 2: Influence of time and level of automation over the difference

Parameter	Time model		Time + Condition model		Time*Condition model	
	$\chi^2(3)$	p-value	$\chi^2(1)$	p-value	$\chi^2(3)$	p-value
Pupil size	1.83	.609	0.40	.528	0.30	.959

Discussion

We studied the impact of automated compared to manual environments on MW and its behavioral and oculometric markers. The automated condition revealed significantly lower TLX scores compared to the manual condition, showing a protocol in line with usual goals regarding automation introduction (Wiener, 1988). Performance remained very high throughout both conditions whereas MW increased in the automated condition, ruling out the possibility that our attentional reports might be significantly influenced by poor performance (Head & Helton, 2016). This demonstrates that our results were as close as possible to ecological settings. Building on this, three main results have been shown: (1) MW increases after some time has elapsed in an automated environment, (2) there is a difference in pupil diameter between MW and focus episodes but not for other oculometric markers and (3) pupillometric difference between attentional states remains stable through time and condition. We discuss these results below.

The first result is the significant increase of the MW frequency under the automated condition between blocks two and three. No significant time-related evolution of MW was observed under the manual condition. Since both conditions lasted the same amount of time, had similar number of actions and pursued the same goal – avoid incoming obstacles –, time-related phenomena (drowsiness, habituation, tiredness) cannot explain entirely the fact that MW increased only under the automated condition. The absence of MW increase in the manual condition is interesting considering the well-established vigilance decrement observed in sustained attention (Cabon et al., 1993; Davies & Parasuraman, 1982; Jeroski, Miller, Langhals, & Tripp, 2014; Mackworth, 1948). It may point to a fundamental difference between MW and vigilance decrement when considering the influence of automation. However mediating factors have still to be investigated, such as anxiety and motivation, which have demonstrated essential link with both phenomena separately (Killingsworth & Gilbert, 2010; Szalma et al., 2004; Szalma & Matthews, 2015). Nevertheless, the level of automation alone cannot explain the observed data. Even though MW frequency highlighted significant differences between conditions, the trend did not evolve linearly with time-on-task and showed no difference between conditions for the first two blocks. Moreover, this evolution happened despite TLX scores remaining low for both conditions, which rules out the possibility that MW may be explained by workload evolution. Together, these findings argue for an effect linked to time spent supervising automation.

There are two explanations, which may be complementary, for this interaction between time and level of automation over MW frequency. First, complacency might be generated by the high reliability of the system and lower monitoring performance. Complacency is an issue of monitoring automation generated by an uncritical reliance on the system (Parasuraman, Molloy, & Singh, 1993). Complacency has been linked to higher reaction time (Bahner, Hüper, & Manzey, 2008; Manzey, Bahner, & Hüper, 2006), loss of situation awareness (Endsley & Kiris, 1995) and failures of detection (Parasuraman et al., 1993). In our experiment, participants encountered no error during the first three blocks. Given that the system never did any miss or error, participants may have thought that it would remain perfectly reliable. In this context, their perception of the required workload might evolve: since the automated system does not seem to require their attention to function properly, participants would redirect their cognitive resources towards more personal matters and mind-wander more. The higher perceived workload under the manual condition supports our analysis. Moreover, this could explain why participants, novice in supervising the system, exhibited an increase of MW frequency only after some time, while pilots in Casner and Schooler (2015), who were pilots with thousands of hours dealing with autopilot, mind wandered immediately without temporal evolution. These evidences suggest a mediating influence of system familiarity in MW frequency temporal evolution. This position would introduce a third possibility within the overload/underload theory debate (Pattyn et al., 2008; Warm, Parasuraman, & Matthews, 2008). Although the task complexity does not change, the operator's perception could evolve based on their trust in the system and their feelings toward the overall situation. As pointed by Seli and colleagues (Seli, Carriere, & Smilek, 2015; Seli et al., 2016), there is strong evidence that people can exert some control over their MW. This follows Casner and Schooler's (2015) results, who demonstrated that cognitive resources freed by automation in peaceful situations are not allocated to task planning, but rather to MW. Our analysis is in line with studies that observed MW increase in a low probability signal environment (Berthié et al., 2015; Casner & Schooler, 2015; Galera et al., 2012), with the time elapsed performing the task (McVay & Kane, 2009; Smallwood, Baracaia, Lowe, & Obonsawin, 2003; Smallwood, Riby, Heim, & Davies, 2006) and the view of complacency as a multiple-task strategy (Bahner et al., 2008; Moray & Inagaki, 2000).

Operators save cognitive resources allocated to the low-event automated task in order to perform better on another task – MW –, which is considered more interesting or useful, independently of experiment instructions.

The second possible explanation is a decoupling of operators' attention from the task at hand. When dealing with automation, operators give up their direct control over the system for a monitoring role in the supervisory control loop (Moray, 1986; Sheridan, 1992). They may experience a loss of agency – i.e., the ability to feel in control (Wegner, 2002). Multiple studies pointed to a limit to the automation level beyond which users felt less in control (Berberian, Sarrazin, Le Blaye, & Haggard, 2012; Coyle, Moore, Kristensson, Fletcher, & Blackwell, 2012), leading to a form of disengagement from the task at hand (Haggard, 2017). Interestingly, Szalma (2014) described a similar disengagement when applying the Self-Determination Theory (Ryan & Deci, 2000) to human-system interactions. The inability of a system to support autonomous behavior may lower motivation and create an externalization of task goals – i.e. a process by which operators rejects the value of a goal. In our experiment, since participants do not validate but rather only acknowledge the system's actions, they could firstly experience a loss of agency, their motivation would decrease, leading to a faint sense of responsibility. This process chain could lead participants to reallocate cognitive resources from the task to MW, unconsciously trying to optimize time and mental resources from their perspective. Further studies are needed to distinguish the respective impacts of agency drop and complacency on MW emergence.

Our second result concerns oculometric measures. We highlighted a lower pupil diameter during MW, as did several studies on MW (Faber, Bixler, & D'Mello, 2017;

Grandchamp et al., 2014; Mittner et al., 2014). Our probes included action required for 8 out of 20 probes in each condition, ruling out the possibility that performance may have significantly influenced subsequent attentional reports (Head & Helton, 2016). Moreover, literature on vigilance already linked a lower pupil baseline to periods of lower sensibility to external stimuli (K. McIntire, P. McIntire, Mckinley, & Goodyear, 2014; Nishiyama, Tanida, Kusumi, & Hirata, 2007). Taken together, these results are in line with the view of MW as a phenomenon inducing a decoupling from the environment. However, other research linked large pupils with slow and inaccurate responses (Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010; Smallwood et al., 2011), or more directly to MW during a word-by-word reading task (Franklin et al., 2013). A recent study by Konishi et al. (2017) was aimed at explaining these results to all appearances contradictory. During 0-back and 1-back tasks, they observed a smaller pupil preceding MW reports. They also linked a higher pupil baseline and slower or inaccurate responses, highlighting a different state of under-processing of external stimuli and ruling out a potential increase in pupil diameter during MW episodes. These results corroborate our study and stress the need to investigate these attentional states. Contrary to pupillometry, other oculometric measures did not exhibit significant sensitivity to MW. However, our experiment is the first to our knowledge to investigate MW influence over blink, saccade frequency and fixation duration in operational settings. Indeed, previous research used most exclusively reading tasks (Bixler & D'Mello, 2014, 2015; Reichle, Reineberg, & Schooler, 2010; D. Smilek et al., 2010; Uzzaman & Joordens, 2011), with the notable exception of meditation (Grandchamp et al., 2014). Our result could point to important task mediators of MW influence over oculometric markers, such as event rate or cognitive demands.

Finally, the last result is the stability of pupillometric markers with respect to automation and time. Cheyne and colleagues (2009) recently proposed the integration of intensity of environment decoupling as a characteristic of MW episodes. They used a Sustained Attention to Response Task (SART, a form of GO/NOGO task; see (Robertson, Manly, Andrade, Baddeley, & Yiend, 1997) to match errors and reaction time evolution with each level of their model. If this model were true, there is little doubt that physiological markers would show some sensibility to intensity of MW. However, no influence over oculometric markers was observed. Several explanations can be proposed. First, our protocol, which differ from previous protocols, may not be able to uncover such a tendency. Second, intensity may not regulate MW impact over pupillometry. Third, there may not be any intensity in MW episodes, each inducing the same environment decoupling. Indeed, the study Cheyne and colleagues (2009) falls under the concerns expressed by Head and Helton (2016), see next paragraph). Further neural studies are necessary to answer this question.

One could argue about the absence of analysis of performances, in order to clarify the relation between MW and stimuli processing. However, our study aimed to explore a different question: the impact of automation on MW occurrence. Addressing both questions would have required modifications to our protocol – add more conflicts, increase the duration, synchronize probes with conflicts –, with the possibility to introduce biases and produce an environment far from an ecological one. In such condition, OOTL phenomenon occurrence would be difficult to induce.

Nevertheless, the question of how MW influences performance remains to be answered. The extended literature on the subject (Esterman, Noonan, Rosenberg, &

DeGutis, 2013; Smallwood et al., 2003; Smilek, Carriere, & Cheyne, 2010a; Thomson, Seli, Besner, & Smilek, 2014) was recently criticized by Head and Helton (2016). They put forward the possibility that poor performance observed before MW may influence subsequent attentional reports, and not the other way around. The result of their reading study did not show any significant link between MW and awareness of stimuli. Certainly, studies using high rates of discrete events without high cognition – like the Sustained Attention to Response Task (SART) – may be particularly biased by this logical flaw, as performance monitoring and self-corrections are easy to perform. On the other hand, continuous metrics – as in tracking tasks (Kam et al., 2012; Yanko & Spalek, 2014) – cannot be similarly biased, as poor performance evaluation is harder and would lead to a direct correction visible in the signal. Similarly, studies measuring stimuli awareness or recognition – like reading – may avoid this flaw, as performance is evaluated either at the end of the session (Franklin et al., 2013), or not at all (Uzzaman & Joordens, 2011). Be that as it may, further research is needed to identify parameters mediating the perceptual decoupling induced by MW.

In the near future, the massive use of automation in everyday systems will reinforce the OOTL phenomenon. Our results show that automation increases MW frequency after some time. The MW literature in ecological tasks already highlighted how the phenomenon increases the risks in critical environments. Such results stress the necessity to study in more detail the relation between MW and the OOTL performance problem. Possible improvements include the study of reliability and complacency by manipulating the number of conflicts and automation errors. Another possibility is to highlight the impact of the operator’s engagement in the task. Finally, perceived workload is not to be overlooked. The use of electroencephalograms would allow continuous measurement to precisely assess its impact over MW frequency. However, such a protocol requires the influence of perceived workload and MW over neural measures to be discriminated. Eventually, the expected outcome is to design automated systems able to adapt themselves to operators' MW episodes. We hope that such a system may enhance safety in critical automated environments.

Acknowledgements

We thank the *Direction Générale de l'Armement* (DGA) for their financial support to the first author. This work was supported by the Agence Nationale de la Recherche (grant ANR-15-CE26-0010-01).

References

1. AC Del Re. (2014). compute.es: Compute Effect Sizes (Version 0.2-4). Retrieved from https://CRAN.R-project.org/package=compute.es
2. Amalberti, R. (1999). Automation in aviation: A human factors perspective. In D. J. Garland, J. Wise, & D. Hopkin (Eds.), *Aviation Human Factors* (pp. 173–192). Hillsdale, New Jersey: Lawrence Erlbaum Associates. Retrieved from http://www.irit.fr/SIGCHI/old/docs/debat/Automation.in.Avia.LEA98%27.doc
3. Bahner, J. E., Hüper, A.-D., & Manzey, D. (2008). Misuse of automated decision aids: Complacency, automation bias and the impact of training experience. *International Journal of Human-Computer Studies*, *66*(9), 688–699. https://doi.org/10.1016/j.ijhcs.2008.06.001

4. Bastian, M., Lerique, S., Adam, V., Franklin, M. S., Schooler, J. W., & Sackur, J. (2017). Language facilitates introspection: Verbal mind-wandering has privileged access to consciousness. *Consciousness and Cognition*, *49*, 86–97. https://doi.org/10.1016/j.concog.2017.01.002
5. Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/doi:10.18637/jss.v067.i01
6. Baxter, G., Rooksby, J., Wang, Y., & Khajeh-Hosseini, A. (2012). The ironies of automation … still going strong at 30? In *Proceedings of ECCE 2012 Conference*. Edinburgh, North Britain.
7. Berberian, B., Sarrazin, J.-C., Le Blaye, P., & Haggard, P. (2012). Automation Technology and Sense of Control: A Window on Human Agency. *PLoS ONE*, *7*(3), e34075. https://doi.org/10.1371/journal.pone.0034075
8. Berthié, G., Lemercier, C., Paubel, P.-V., Cour, M., Fort, A., Galera, C., … Maury, B. (2015). The restless mind while driving: drivers' thoughts behind the wheel. *Accident Analysis & Prevention*, *76*, 159–165. https://doi.org/10.1016/j.aap.2015.01.005
9. Bixler, R., & D'Mello, S. (2014). Toward Fully Automated Person-Independent Detection of Mind Wandering. In *User Modeling, Adaptation, and Personalization* (pp. 37–48). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-08786-3_4
10. Bixler, R., & D'Mello, S. (2015). Automatic gaze-based user-independent detection of mind wandering during computerized reading. In *Proceedings of the*

23rd User Modeling and User-Adapted Interactions conference (pp. 31–43). Dublin, Ireland. https://doi.org/10.1007/s11257-015-9167-1

11. Cabon, P., Coblentz, A., Mollard, R., & Fouillot, J. P. (1993). Human vigilance in railway and long-haul flight operation. *Ergonomics*, *36*(9), 1019–1033. https://doi.org/10.1080/00140139308967974
12. Casner, S. M., & Schooler, J. W. (2015). Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. *Consciousness and Cognition*, *35*, 33–41. https://doi.org/10.1016/j.concog.2015.04.019
13. Cegarra, J., & Morgado, N. (2009). Étude des propriétés de la version francophone du NASATLX. In B. Cahour, F. Anceaux, & A. Giboin (Eds.), *Acte du 5ème colloque de psychologie ergonomique (EPIQUE'2009)* (pp. 233–239). Nice, France.
14. Cheyne, A. J., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, *111*(1), 98–113. https://doi.org/10.1016/j.cognition.2008.12.009
15. Coyle, D., Moore, J., Kristensson, P. O., Fletcher, P., & Blackwell, A. (2012). I did that! Measuring users' experience of agency in their own actions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2025–2034). Austin, TX: ACM. https://doi.org/10.1145/1518701.1518766
16. Cummings, M. L., Sasangohar, F., Thornburg, K. M., Xing, J., & D'Agostino, A. (2010). Human-system interface complexity and opacity part i: literature review.

Massachusetts Institute of Technology, Cambridge, MA. Retrieved from https://www.researchgate.net/profile/Farzan_Sasangohar/publication/263315585_Human-system_interface_complexity_and_opacity/links/0deec53accc4413fb2000000.pdf

17. Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. London, England: Academic Press.
18. Dündar, C. (2015). *The effects of mind wandering on simulated driving performance*. Retrieved from http://etd.lib.metu.edu.tr/upload/12618921/index.pdf
19. Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The Unengaged Mind: Defining Boredom in Terms of Attention. *Perspectives on Psychological Science*, *7*(5), 482–495. https://doi.org/10.1177/1745691612456044
20. Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(2), 381–394.
21. Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2013). In the Zone or Zoning Out? Tracking Behavioral and Neural Fluctuations During Sustained Attention. *Cerebral Cortex*, *23*(11), 2712–2723. https://doi.org/10.1093/cercor/bhs261
22. Faber, M., Bixler, R., & D’Mello, S. K. (2017). An automated behavioral measure of mind wandering during computerized reading. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0857-y

23. Franklin, M. S., Broadway, J. M., Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2013). Window to the wandering mind: Pupillometry of spontaneous thought while reading. *The Quarterly Journal of Experimental Psychology*, *66*(12), 2289–2294. https://doi.org/10.1080/17470218.2013.858170
24. Fraulini, N. W., Hancock, G. M., Neigel, A. R., Claypoole, V. L., & Szalma, J. L. (2017). A critical examination of the research and theoretical underpinnings discussed in Thomson, Besner, and Smilek (2016). *Psychological Review*, *124*(4), 525–531. https://doi.org/10.1037/rev0000066
25. Fried, R., Schettlinger, K., & Borowski, M. (2014). robfilter: Robust Time Series Filters (Version 4.1). Retrieved from https://CRAN.R-project.org/package=robfilter
26. Galera, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribereau-Gayon, R., … Lagarde, E. (2012). Mind wandering and driving: responsibility case-control study. *British Medical Journal*, *345*(dec13 8), e8105–e8105. https://doi.org/10.1136/bmj.e8105
27. Gerbert, K., & Kemmler, R. (1986). The causes of causes: determinants and background variables of human factor incidents and accidents. *Ergonomics*, *29*(11), 1439–1453. https://doi.org/10.1080/00140138608967257
28. Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252–269. https://doi.org/10.3758/CABN.10.2.252

29. Golchert, J., Smallwood, J., Jefferies, E., Seli, P., Huntenburg, J. M., Liem, F., … Margulies, D. S. (2016). Individual variation in intentionality in the mind-wandering state is reflected in the integration of the default-mode, fronto-parietal, and limbic networks. *NeuroImage*, *146*, 226–235. https://doi.org/10.1016/j.neuroimage.2016.11.025
30. Gouraud, J., Delorme, A., & Berberian, B. (2017). Autopilot, Mind Wandering, and the Out of the Loop Performance Problem. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00541
31. Grandchamp, R., Braboszcz, C., & Delorme, A. (2014). Oculometric variations during mind wandering. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00031
32. Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews Neuroscience*, *18*(4), 196–207. https://doi.org/10.1038/nrn.2017.14
33. Hancock, P. A. (2013). In search of vigilance: The problem of iatrogenically created psychological phenomena. *American Psychologist*, *68*(2), 97–109. https://doi.org/10.1037/a0030214
34. Hart, S. G., & Staveland, L. E. (1988). Development of NASA- TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139–203.
35. Head, J., & Helton, W. S. (2016). The troubling science of neurophenomenology. *Experimental Brain Research*. https://doi.org/10.1007/s00221-016-4623-7
36. 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), 17–21. https://doi.org/10.1177/0018720810391530

37. Hothorn, T., Bretz, F., Westfall, P., Herberger M., R., Schuetzenmeister, A., & Schreibe, S. (2017). multComp: Simultaneous Inference in General Parametric Models (Version 1.4-8). Retrieved from https://CRAN.R-project.org/package=multcomp
38. Jeroski, J., Miller, M. E., Langhals, B., & Tripp, L. (2014). Impact of Vigilance Decrement upon Physiology Measures. In *IIE Annual Conference. Proceedings* (p. 1184). Institute of Industrial Engineers-Publisher. Retrieved from http://www.xcdsystem.com/iie2014/abstract/finalpapers/I471.pdf
39. Kam, J. W. Y., Dao, E., Blinn, P., Krigolson, O. E., Boyd, L. A., & Handy, T. C. (2012). Mind wandering and motor control: off-task thinking disrupts the online adjustment of behavior. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00329
40. Kam, J. W. Y., Dao, E., Farley, J., Fitzpatrick, K., Smallwood, J., Schooler, J. W., & Handy, T. C. (2011). Slow fluctuations in attentional control of sensory cortex. *Journal of Cognitive Neuroscience*, *23*(2), 460–470.
41. Killingsworth, M. A., & Gilbert, D. T. (2010). A Wandering Mind Is an Unhappy Mind. *Science*, *330*(6006), 932–932. https://doi.org/10.1126/science.1192439
42. K. McIntire, L., P. McIntire, J., Mckinley, R., & Goodyear, C. (2014). Detection of vigilance performance with pupillometry. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 167–174). Safety Harbor, FL, USA: ACM.

43. Konishi, M., Brown, K., Battaglini, L., & Smallwood, J. (2017). When attention wanders: Pupillometric signatures of fluctuations in external attention. *Cognition*, *168*, 16–26. https://doi.org/10.1016/j.cognition.2017.06.006

44. Lawrence, M. A. (2016). ez: Easy Analysis and Visualization of Factorial Experiments (Version 4.4-0). Retrieved from https://CRAN.R-project.org/package=ez

45. Lemercier, A. (2014). *Développement de la pupillométrie pour la mesure objective des émotions dans le contexte de la consommation alimentaire* (Psychologie). Université Paris 8.

46. Lerner, N., Baldwin, C., Higgins, J. S., Lee, J., & Schooler, J. (2015). Mind Wandering While Driving: What Does it Mean and What do we do about it? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *59*(1), 1686–1690. https://doi.org/10.1177/1541931215591364

47. Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quaterly Journal of Experimental Psychology*, *1*, 6–14. https://doi.org/10.1080/17470214808416738

48. Mair, P., Schoenbrodt, F., & Wilcox, R. (2017). WRS2: Wilcox robust estimation and testing.

49. Manzey, D., Bahner, J. E., & Hüper, A.-D. (2006). Misuse of automated aids in process control: Complacency, automation bias and possible training interventions. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 50, pp. 220–224). Los Angeles, CA: Sage Publications.

Retrieved from http://journals.sagepub.com/doi/abs/10.1177/154193120605000303

50. McMillan, R. L., Kaufman, S. B., & Singer, J. L. (2013). Ode to positive constructive daydreaming. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00626
51. McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 196.
52. Methot, L. L., & Huitema, B. E. (1998). Effects of signal probability on individual differences in vigilance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *40*(1), 102–110.
53. Mittner, M., Boekel, W., Tucker, A. M., Turner, B. M., Heathcote, A., & Forstmann, B. U. (2014). When the Brain Takes a Break: A Model-Based Analysis of Mind Wandering. *Journal of Neuroscience*, *34*(49), 16286–16295. https://doi.org/10.1523/JNEUROSCI.2062-14.2014
54. Moray, N. (1986). Monitoring behavior and supervisory control. In K. Boff (Ed.), *Handbook of perception and human performance* (Vol. 2, pp. 1–51). New York, NY: Wiley.
55. Moray, N., & Inagaki, T. (2000). Attention and complacency. *Theoretical Issues in Ergonomics Science*, *1*(4), 354–365. https://doi.org/10.1080/14639220052399159

56. Mosier, K. L., Skitka, L. J., & Korte, K. J. (1994). Cognitive and social psychological issues in flight crew/automation interaction. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: current research and trends* (pp. 191–197). Cambridge: University Press.
57. Naujoks, F., Purucker, C., & Neukum, A. (2016). Secondary task engagement and vehicle automation – Comparing the effects of different automation levels in an on-road experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*, *38*, 67–82. https://doi.org/10.1016/j.trf.2016.01.011
58. Nishiyama, J., Tanida, K., Kusumi, M., & Hirata, Y. (2007). The pupil as a possible premonitor of drowsiness. In *29th Annual International Conference of the IEEE* (pp. 1586–1589). Lyon, France: IEEE. https://doi.org/10.1016/j.neuron.2015.09.012
59. Parasuraman, R. (1987). Human-computer monitoring. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *29*(6), 695–706.
60. Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance Consequences of Automation-Induced “Complacency.” *The International Journal of Aviation Psychology*, *3*(1), 1–23.
61. Pattyn, N., Neyt, X., Henderickx, D., & Soetens, E. (2008). Psychophysiological investigation of vigilance decrement: Boredom or cognitive fatigue? *Physiology & Behavior*, *93*(1-2), 369–378. https://doi.org/10.1016/j.physbeh.2007.09.016
62. R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

63. Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye Movements During Mindless Reading. *Psychological Science*, *21*(9), 1300–1310. https://doi.org/10.1177/0956797610378686
64. Revelle, W. (2017). psych: Procedures for Personality and Psychological Research (Version 1.7.5). Evanston, Illinois, USA: Northwestern University. Retrieved from https://CRAN.R-project.org/package=psych
65. 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.
66. RStudio Team. (2015). RStudio: Integrated Development for R (Version 1.0.143). Boston, MA, USA: RStudio Inc. Retrieved from http://www.rstudio.com/
67. Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.
68. Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. *Handbook of Human Factors and Ergonomics*, *2*, 1926–1943.
69. Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *5*, 319–326. https://doi.org/10.1016/j.tics.2011.05.006
70. Seli, P., Carriere, J. S. A., & Smilek, D. (2015). Not all mind wandering is created equal: dissociating deliberate from spontaneous mind wandering. *Psychological Research*, *79*(5), 750–758. https://doi.org/10.1007/s00426-014-0617-x

71. Seli, P., Risko, E. F., & Smilek, D. (2016). On the necessity of distinguishing between unintentional and intentional mind wandering. *Psychological Science*, *27*(5), 685–691. https://doi.org/10.1177/0956797616634068
72. Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. Cambridge: MIT Press.
73. Smallwood, J. (2013). Distinguishing how from why the mind wanders: A process–occurrence framework for self-generated mental activity. *Psychological Bulletin*, *139*(3), 519–535. https://doi.org/10.1037/a0030010
74. Smallwood, J., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition*, *12*(3), 452–484. https://doi.org/10.1016/S1053-8100(03)00018-7
75. Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., … Schooler, J. W. (2011). Pupillometric Evidence for the Decoupling of Attention from Perceptual Input during Offline Thought. *PLoS ONE*, *6*(3), e18298. https://doi.org/10.1371/journal.pone.0018298
76. Smallwood, J., Riby, L., Heim, D., & Davies, J. B. (2006). Encoding during the attentional lapse: Accuracy of encoding during the semantic sustained attention to response task. *Consciousness and Cognition*, *15*(1), 218–231. https://doi.org/10.1016/j.concog.2005.03.003
77. Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958. https://doi.org/10.1037/0033-2909.132.6.946
78. Smallwood, J., & Schooler, J. W. (2015). The Science of Mind Wandering: Empirically Navigating the Stream of Consciousness. *Annual Review of*

Psychology, *66*(1), 487–518. https://doi.org/10.1146/annurev-psych-010814-015331

79. Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010a). Failures of sustained attention in life, lab, and brain: Ecological validity of the SART. *Neuropsychologia*, *48*(9), 2564–2570. https://doi.org/10.1016/j.neuropsychologia.2010.05.002
80. Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010b). Out of Mind, Out of Sight: Eye Blinking as Indicator and Embodiment of Mind Wandering. *Psychological Science*, *21*(6), 786–789. https://doi.org/10.1177/0956797610368063
81. Szalma, J. L. (2014). On the Application of Motivation Theory to Human Factors/Ergonomics: Motivational Design Principles for Human–Technology Interaction. *Human Factors*, *56*(8), pp. 1453–1471. https://doi.org/10.1177/0018720814553471
82. Szalma, J. L., & Matthews, G. (2015). Motivation and Emotion in Sustained Attention.
83. Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier, A., & Eggemeier, F. T. (2004). Effects of sensory modality and task duration on performance, workload, and stress in sustained attention. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(2), 219–233.
84. Thomson, D. R., Besner, D., & Smilek, D. (2016). A critical examination of the evidence for sensitivity loss in modern vigilance tasks. *Psychological Review*, *123*(1), 70–83. https://doi.org/10.1037/rev0000021

85. Thomson, D. R., Seli, P., Besner, D., & Smilek, D. (2014). On the link between mind wandering and task performance over time. *Consciousness and Cognition*, *27*, 14–26. https://doi.org/10.1016/j.concog.2014.04.001
86. Uzzaman, S., & Joordens, S. (2011). The eyes know what you are thinking: Eye movements as an objective measure of mind wandering. *Consciousness and Cognition*, *20*(4), 1882–1886. https://doi.org/10.1016/j.concog.2011.09.010
87. Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance Requires Hard Mental Work and Is Stressful. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152
88. Wegner, D. M. (2002). *The Illusion of Conscious Will* (MA: MIT Press). Cambridge.
89. Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12).
90. Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis. *Springer-Verlag New York*.
91. Wiener, E. L. (1988). *Cockpit automation*. National Aeronautics and Space Administration. Retrieved from http://ntrs.nasa.gov/search.jsp?R=19890047073
92. Wise, J. A., Tilden, D. S., Abbott, D., Dyck, J., & Guide, P. (1994). Managing automation in the cockpit. In *Proceedings of the 24th International Conference of the International Federation of Airworthiness*. Lisbon, Portugal. Retrieved from http://www.faa.gov/training_testing/training/media/cfit/volume2/pdf/pages/page5_07.pdf

93. Yanko, M. R., & Spalek, T. M. (2014). 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. https://doi.org/10.1177/0018720813495280

Figure 1

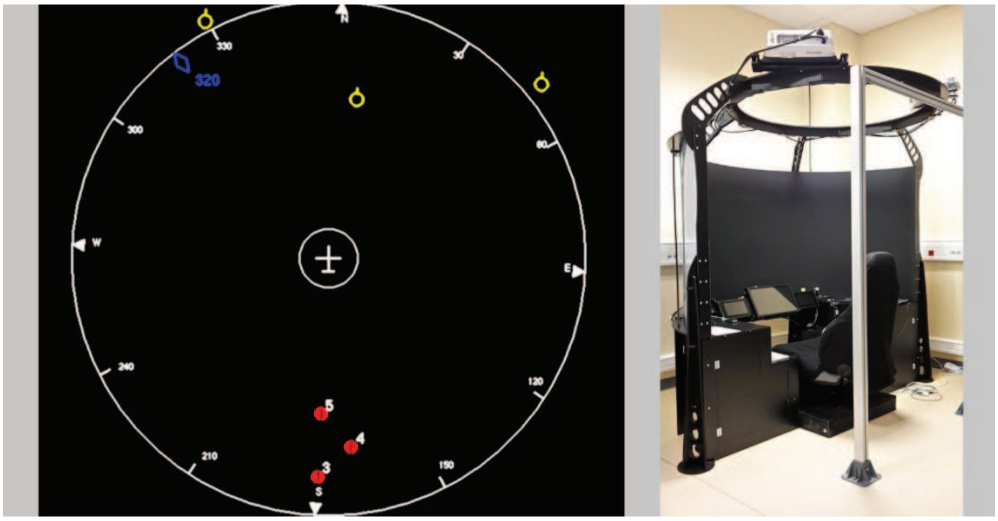
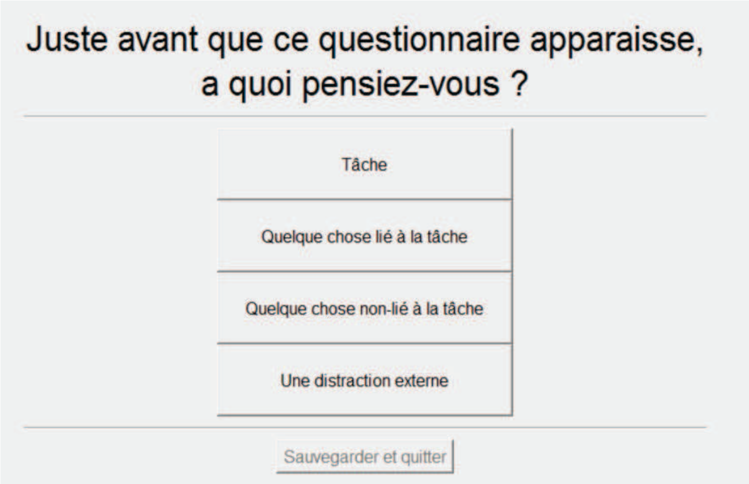


Figure 1. Screenshot of the LIPS interface and the environment. One of the screen is used for the task and the other one for questionnaire probes. For the task, the plane in the center is static and the surround (yellow and red numbered symbols) are moving. During, left and right avoidance maneuvers, again the plane is static and the background is rotated.

Figure 2



Juste avant que ce questionnaire apparaisse, a quoi pensiez-vous ?

Figure 2. Screenshot of the French MW probes. The question is “When this probe appeared, where was your attention directed?” Answers could be “On the task”, “Something related to the task”, “Something unrelated to the task” or “External distraction”

Figure 3

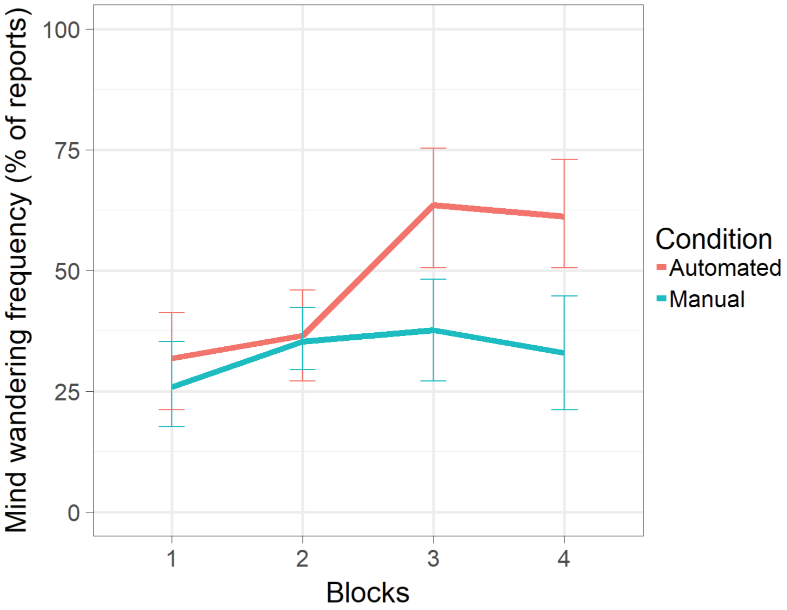


Figure 3: MW frequency evolution according to the condition. Error bars show the 95% confidence intervals based on bootstrap

Figure 4

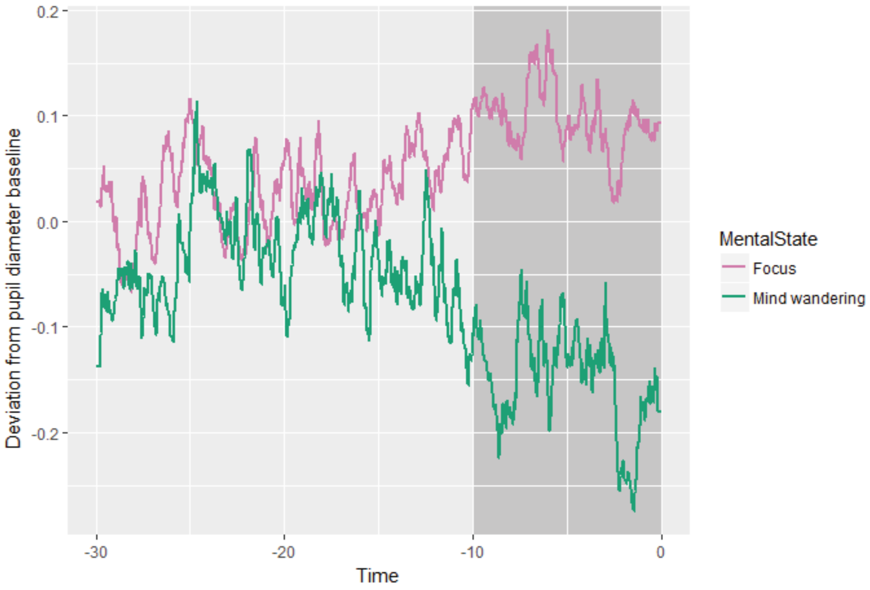


Figure 4: Normalized pupil diameter. Evolution during the 30-second interval preceding probes display – the grey part of the signal is used for computation

Table 1

Table 1: Comparison of oculometric measures during MW and focus episodes

Parameter	**MW values**		**Focus values**		**Mental State model**	
	M	*SD*	*M*	*SD*	$\chi^2(1)$	p-value
Pupil size (mm)	4.90	0.97	5.05	0.97	**7259**	**<.001**
Saccade frequency (sacc/s)	3.92	2.36	3.89	2.39	0.07	.795
Mean fixation duration (s)	$2.87*10^{-1}$	$6.65*10^{-1}$	$3.22*10^{-1}$	$6.52*10^{-1}$	0.08	.774
Blink frequency (blink/s)	$6.90*10^{-2}$	$1.10*10^{-1}$	$5.43*10^{-2}$	$9.81*10^{-2}$	2.09	.148

Table 2

Table 2: Influence of time and level of automation over the difference

Parameter	**Time model**		**Time + Condition model**		**Time*Condition model**	
	$\chi^2(3)$	p-value	$\chi^2(1)$	p-value	$\chi^2(3)$	p-value
Pupil size	1.83	.609	0.40	.528	0.30	.959