



HAL
open science

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.

1 Influence of Automation over Mind Wandering Frequency in Sustained Attention

2 Jonas Gouraud¹, Arnaud Delorme^{2,3}, Bruno Berberian¹

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

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

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

6 California, USA

7

8

9

10 Correspondence concerning this article should be addressed to Jonas Gouraud,

11 ONERA Salon de Provence, 13661, Salon Air, France. Contact: contact@jonasgouraud.com

13

Abstract

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

26 *Keywords:* mind wandering; automation; vigilance; oculometry; complacency;
27 decoupling

28

29

Introduction

30 In order to continuously improve system safety, the critical systems industry
31 makes extensive use of automation (Baxter, Rooksby, Wang, & Khajeh-Hosseini, 2012;
32 Parasuraman, 1987). In cockpits (Wise, Tilden, Abbott, Dyck, & Guide, 1994), in cars
33 (Naujoks, Purucker, & Neukum, 2016), and in power plant consoles (Cummings,
34 Sasangohar, Thornburg, Xing, & D'Agostino, 2010), automation has been introduced to
35 increase performance and respond to new safety requirements. Unfortunately, while
36 implementing higher levels of automation indeed improves the efficiency and capacity of
37 a system, it also creates new challenges for human operators. Particularly, the externally
38 imposed task to maintain sustained attention focused for long periods of time in low
39 probability environments causes progressive vigilance decrement – or invigilance
40 increment (Hancock, 2013) – preventing efficient automation supervising (Amalberti,
41 1999). As targets are hidden – naturally, voluntarily or because of poor display design –,
42 the task to detect and react to these targets is often stressful and increasingly difficult
43 (Mackworth, 1948). These problems result in out-of-the-loop (OOTL) performance
44 problem, referring to a performance decrease whenever attempts are made to regain
45 manual control after a critical system failure.

46 Such problems have been studied in laboratories (Endsley & Kiris, 1995; Sarter,
47 Woods, & Billings, 1997), but are also regularly reported in operational conditions.
48 Mosier and collaborators (1994) examined NASA's Aviation Safety Reporting System
49 (ASRS) database and reported that 77% of the incidents involved an over-reliance on
50 automation leading to a probable vigilance failure. Similarly, Gerbert and Kemmler
51 (1986) studied German aviators' anonymous responses to questionnaires about

52 automation-related incidents and pointed out failures of vigilance as the largest
53 contributor to human error. Several studies showed that efficient sustained attention for
54 hours cannot be achieved (Caban, Coblentz, Mollard, & Fouillot, 1993; Mackworth,
55 1948; Methot & Huitema, 1998).

56 Such a context may favor the occurrence of mind wandering (MW) episodes. MW
57 is a family of experiences relating to the mind's tendency to engage in thoughts unrelated
58 to the here and now (Smallwood & Schooler, 2006). It is an ubiquitous phenomenon that
59 can be intentional or spontaneous (Golchert et al., 2016; Seli, Risko, & Smilek, 2016), be
60 guided or unguided (Smallwood, 2013), emerge when performing a task or at rest
61 (Smallwood, Baracaia, Lowe, & Obonsawin, 2003) while its ignition point can be
62 triggered by the environment or generated internally (McMillan, Kaufman, & Singer,
63 2013; Smallwood & Schooler, 2006). In the following paper, we focused on MW when
64 performing a task without discriminating other dimensions. MW is more likely to occur
65 in monotonous environments (Eastwood, Frischen, Fenske, & Smilek, 2012), or when
66 operators perform familiar (Bastian et al., 2017) or long tasks (Smallwood & Schooler,
67 2015). Its occurrence favors a decoupling from the ongoing task at perceptual and stimuli
68 processing levels (Kam et al., 2012; Schooler et al., 2011), which can be seen both on
69 behavioral and physiological data. Reading tasks were particularly used to uncover the
70 influence of MW over oculometric markers like blink frequency (Smilek, Carriere, &
71 Cheyne, 2010b), fixation duration and saccade frequency (Uzzaman & Joordens, 2011).
72 In simulators, Yanko and Spalek (2014) studied MW influence over driving performance.
73 They observed a longer reaction time to unexpected events, a shorter headway distance
74 and a higher velocity. Their results were corroborated by other studies in driving

75 environments (Dündar, 2015; He, Becic, Lee, & McCarley, 2011; Lerner, Baldwin,
76 Higgins, Lee, & Schooler, 2015).

77 Given that MW diverts an operator's attention from his primary task, it could play
78 an important role in vigilance failures observed in highly reliable automated
79 environments. Casner and Schooler (2015) studied the impact of automation on MW in
80 an aeronautical context. Their results on 16-minute sessions did not show a significant
81 correlation between automation and the frequency of MW reports. However, the
82 propensity to mind wander appeared to increase when everything seemed under control.
83 Supervising ultra-reliable systems could encourage operators to decrease cognitive
84 resources allocated to the monitoring task. In that context, time saved by automation,
85 which should normally be used for other productive tasks and for monitoring, could
86 instead be filled by task-unrelated thoughts. Operators in such a state would not be
87 prepared to regain manual control over the system in response to rare critical events.
88 Such analysis is already considered in the debate regarding the origin of the vigilance
89 decrement (Fraulini, Hancock, Neigel, Claypoole, & Szalma, 2017; Pattyn, Neyt,
90 Henderickx, & Soetens, 2008; Thomson, Besner, & Smilek, 2016), recent evidences
91 showing that both phenomena share many features (Gouraud, Delorme, & Berberian,
92 2017).

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

97

98

Material and methods

99

Participants

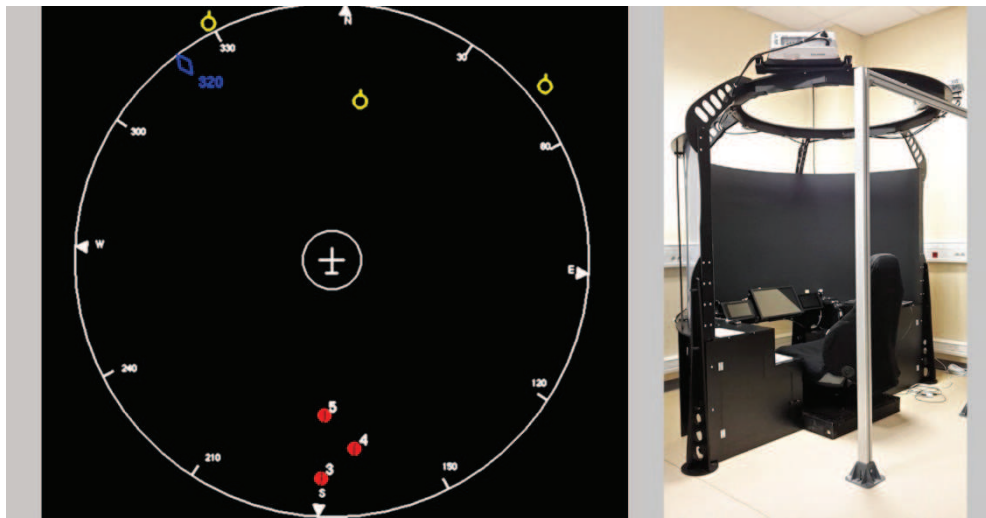
100 17 participants (5 female) performed the experiment (age ranging from 21 to 42
101 years old; $M = 27.3$, $SD = 6.0$). The participants enrolled in this study were volunteers
102 from our company (ONERA organization). All participants had normal or corrected-to-
103 normal visual acuity. All participants signed a written declaration of informed consent.
104 The protocol was conducted in accordance with the Declaration of Helsinki.

105

Task

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

116



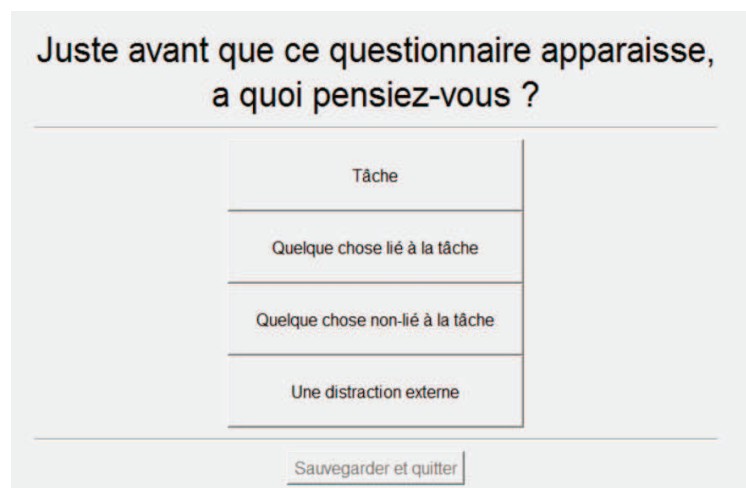
117

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

122

123 **MW probes.** Python 3.6 was used to program mental probes. On average every 2
124 minutes, the probe appeared on a secondary 10-inch screen next to the main screen. For
125 technical reasons, the obstacle avoidance task was not paused when the probe was
126 displayed. Participants were asked to fill it as soon as it appeared, and any successful or
127 failed trial during this interval would not be taken into account. Participants were
128 informed that the probe was not part of the evaluation to lower the impact of instructions
129 over their natural propensity to mind wander. Participants were required to answer the
130 following questions (originally in French, see Figure 2): “When this probe appeared,
131 where was your attention directed?” Answers could be “On the task” (e.g., thinking about
132 the next obstacle, the decision to make, the incoming waypoint), “Something related to
133 the task” (e.g., thinking about performance, interface items, last trial), “Something
134 unrelated to the task” (e.g., thinking about a memory, their last meal, or a body sensation,

135 hereafter defined as MW) or “External distraction” (e.g., conversation, noise). The
136 preceding examples were given to participants to illustrate each category. We were
137 primarily interested in reports of being “On the task” and MW reports. Reports of
138 thoughts “Related to the task” were integrated to avoid participants to report MW when
139 thinking about their performance (Head & Helton, 2016). Noises were integrated to avoid
140 participants to report MW if they were focused on any external signal.
141



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

Tâche
Quelque chose lié à la tâche
Quelque chose non-lié à la tâche
Une distraction externe

Sauvegarder et quitter

142

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

146

147 **Conditions.** Two conditions were proposed. The first one was the “manual”
148 condition and required participants to manually avoid obstacles. The system detected
149 obstacles on the trajectory 13s before impact. Then, an orange circle appeared around the
150 UAV and the participant could initiate an avoidance maneuver. Participants were able to
151 choose the way in which they wished to avoid the obstacle by clicking on “*Evitement*
152 *Gauche*” (left maneuver) or “*Evitement Droite*” (right maneuver). Once they clicked, the

153 simulator turned the trajectory of the UAV on the chosen side, following a predefined
154 angle. Each obstacle had a safe circle similar to that of the UAV (see Figure 1). A
155 collision warning – i.e., an orange circle around both the UAV and the obstacle with a
156 message “Collision” – was displayed if the UAV safe circle penetrated inside the obstacle
157 safe circle. A trial with a collision warning triggered was marked as failed. To resume the
158 initial trajectory, they had to click on the “*Retour trajectoire*” (return to original
159 trajectory) button. If no action was taken within 16 seconds after the first change in
160 trajectory, the aircraft automatically resumed the trajectory and the trial was marked as
161 failed.

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

174 **Procedure.** Participants were explicitly instructed that detection accuracy was
175 more important than speed in button clicks. Each participant performed the two

176 conditions on two separate days in a counterbalanced way. Each day started with an
177 explanation of the task, followed by a 10-minute training period and a 45-minutes session
178 under the proper condition. Each session contained 60 clusters of obstacles. Each cluster
179 was considered a trial. They were separated by 45 seconds on average. 20 probes were
180 answered under each condition. The distribution of probes was not correlated with events
181 on the obstacle avoidance task in order to avoid performance to influence MW reports
182 (Head & Helton, 2016). The automated condition included 8 conflicts with a probe within
183 the 10-seconds interval following the conflict. Participants encountered one system error
184 (where they had to click on the “*Changement d’altitude*” button) during training, and
185 another during the automated condition at the end of the third block. Under the manual
186 condition, participants encountered at the end of the third block a conflict impossible to
187 avoid. Both the automation error and this conflict were not followed by an attentional
188 probe for at least 10 seconds after.

189 **Data recording**

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

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

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

202 **Oculometry.** Oculometric data was recorded using the hardware SmartEye Pro
203 3.0 and the software SmartEye 6.2.4. The system included 2 infrared illuminators and
204 3 cameras (120Hz) placed above the screen to avoid any direct contact with the
205 participant (see Figure 1). Gaze calibration was performed using a 4-point grid.

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

209 **Data Analysis**

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

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

221 We performed filtering on pupillometry using the R package reshape (Wickham,
222 2007), psych (Revelle, 2017), ggplot2 (Wickham, 2009, p. 2) and robfilter (Fried,
223 Schettlinger, & Borowski, 2014). Filtering was done in two passes, following a method
224 already used in the literature (Grandchamp, Braboszcz, & Delorme, 2014). Firstly, we
225 filtered the signal. Pupil diameter had to be between 1 and 10 mm (due to the physical
226 limits of pupil diameter, see (Lemercier, 2014), had to be less than 80% different from
227 the preceding value (due to pupil dynamic limits) and had to be of a quality (computed by
228 the SmartEye software) over 0.01. Extracts were discarded if their resulting pupil
229 diameter series consisted of less than 70% compliant values. The proportion of extracts
230 excluded due to low quality (9.6%) is in line with that excluded in other investigations
231 (Smallwood et al., 2011). Resulting extracts were completed using basic linear
232 interpolation. A second filtering pass was applied with a median filter (moving window
233 of 50 frames). Finally, the data of each participant were normalized by subtracting the
234 mean and dividing by the root mean square of all good-enough quality extracts for this
235 participant.

236 Fixations, saccades and blinks were computed by the SmartEye Pro software.
237 Blinks were computed using sliding windows of 700ms. Saccades were defined in
238 SmartEye Pro parameters as gaze velocity over 35 deg/s. Saccades were limited to
239 200ms. Fixations were frames where the gaze velocity remained below 15 deg/s.

240 **Performance.** Performance was assessed by determining if participants clicked
241 when they were required to do so. Reaction time were computed by comparing
242 participants button click time delay in the manual condition to the moment the system

243 detected an obstacle, and in the automated condition to the time at which they
244 acknowledged each automation decision.

245

246

Results

247

Mind Wandering Frequency

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

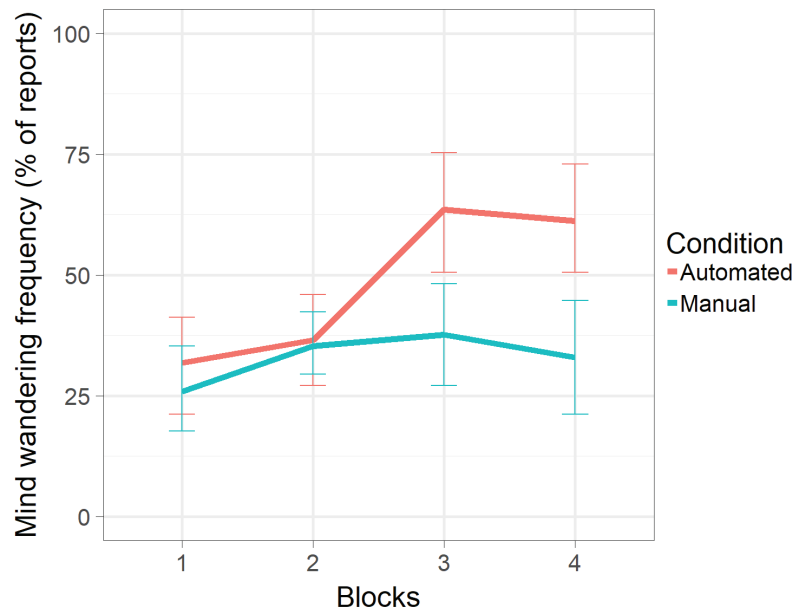
264

265

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 \times 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

266 et al., 2017) and *mes* (AC Del Re, 2014) functions. For the manual condition, all
 267 differences were non-significant ($p > .366$). For the automated condition, the third and
 268 fourth blocks had significantly higher MW frequency compared to the first block, $p =$
 269 $.001$, $d = 0.54$ and $p = .003$, $d = 0.32$, respectively. Similarly, the blocks 3 and 4 had
 270 significantly higher MW frequency compared to the block 2, $p = .007$, $d = 0.12$ and $p =$
 271 $.016$, $d = 0.12$, respectively.
 272



273

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

276

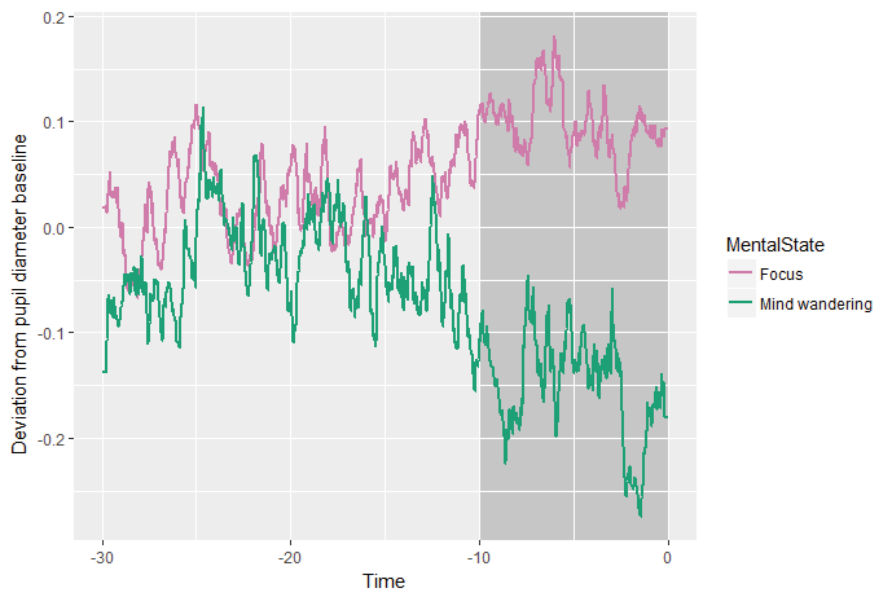
277 NASA TLX Scores

278 Each participant filled 2 TLX questionnaires (one after each session). The mean
 279 score for each TLX of each subject (see Figure 4) varied substantially (ranging from 2 to
 280 14.17, $M = 5.81$, $SD = 2.44$). Shapiro-Wilk's test indicated that the assumption of
 281 normality had been violated for the TLX values, $W = .921$, $p = .012$. Therefore, we used

282 Wilcox's robust version of the t -test proposed in the WRS2 package (Mair, Schoenbrodt,
 283 & Wilcox, 2017). On average, participants perceived that the automated ($M = 4.93$, $SD =$
 284 $.50$) condition required more cognitive resources than the manual ($M = 6.68$, $SD = .61$)
 285 condition, $t(10) = -3.35$, $p = .007$, $d = 0.78$. TLX scores show that our automated
 286 condition succeeded in lowering workload.

287

288



289

290

Figure 4: Normalized pupil diameter. Evolution during the 30-second interval preceding probes

291

display – the grey part of the signal is used for computation

292

293

Oculometry

294

Influence of MW over oculometric measures. Oculometric measures were first

295

analyzed using the 10 seconds preceding each probe. We used the lmer function (Bates,

296

Maechler, Bolker, & Walker, 2015, p. 4) to perform a linear mixed-effect analysis despite

297

missing values – “external distraction” reports and bad quality extracts excluded. As

298 random effects, we had intercepts for subjects. Visual inspection of residual plots did not
 299 reveal any obvious deviations from normality or homoscedasticity. P-values were
 300 obtained by likelihood ratio tests using ANOVA on the full model against the models
 301 with no fixed effect, with only block, with block and condition, and with block and
 302 condition and their interaction. The results are shown in Table 1. On average, participants
 303 showed a significantly smaller pupil during MW episodes (see Figure 4). There was no
 304 main effect of MW on other markers.

305

306

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

Parameter	MW values		Focus values		Mental State model	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	$\chi^2(1)$	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)	$2.87 \cdot 10^{-1}$	$6.65 \cdot 10^{-1}$	$3.22 \cdot 10^{-1}$	$6.52 \cdot 10^{-1}$	0.08	.774
Blink frequency (blink/s)	$6.90 \cdot 10^{-2}$	$1.10 \cdot 10^{-1}$	$5.43 \cdot 10^{-2}$	$9.81 \cdot 10^{-2}$	2.09	.148

307

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

317

318

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

319

320

Discussion

321

We studied the impact of automated compared to manual environments on MW

322

and its behavioral and oculometric markers. The automated condition revealed

323

significantly lower TLX scores compared to the manual condition, showing a protocol in

324

line with usual goals regarding automation introduction (Wiener, 1988). Performance

325

remained very high throughout both conditions whereas MW increased in the automated

326

condition, ruling out the possibility that our attentional reports might be significantly

327

influenced by poor performance (Head & Helton, 2016). This demonstrates that our

328

results were as close as possible to ecological settings. Building on this, three main

329

results have been shown: (1) MW increases after some time has elapsed in an automated

330

environment, (2) there is a difference in pupil diameter between MW and focus episodes

331

but not for other oculometric markers and (3) pupillometric difference between

332

attentional states remains stable through time and condition. We discuss these results

333

below.

334

The first result is the significant increase of the MW frequency under the

335

automated condition between blocks two and three. No significant time-related evolution

336

of MW was observed under the manual condition. Since both conditions lasted the same

337

amount of time, had similar number of actions and pursued the same goal – avoid

338

incoming obstacles –, time-related phenomena (drowsiness, habituation, tiredness) cannot

339 explain entirely the fact that MW increased only under the automated condition. The
340 absence of MW increase in the manual condition is interesting considering the well-
341 established vigilance decrement observed in sustained attention (Caban et al., 1993;
342 Davies & Parasuraman, 1982; Jeroski, Miller, Langhals, & Tripp, 2014; Mackworth,
343 1948). It may point to a fundamental difference between MW and vigilance decrement
344 when considering the influence of automation. However mediating factors have still to be
345 investigated, such as anxiety and motivation, which have demonstrated essential link with
346 both phenomena separately (Killingsworth & Gilbert, 2010; Szalma et al., 2004; Szalma
347 & Matthews, 2015). Nevertheless, the level of automation alone cannot explain the
348 observed data. Even though MW frequency highlighted significant differences between
349 conditions, the trend did not evolve linearly with time-on-task and showed no difference
350 between conditions for the first two blocks. Moreover, this evolution happened despite
351 TLX scores remaining low for both conditions, which rules out the possibility that MW
352 may be explained by workload evolution. Together, these findings argue for an effect
353 linked to time spent supervising automation.

354 There are two explanations, which may be complementary, for this interaction
355 between time and level of automation over MW frequency. First, complacency might be
356 generated by the high reliability of the system and lower monitoring performance.
357 Complacency is an issue of monitoring automation generated by an uncritical reliance on
358 the system (Parasuraman, Molloy, & Singh, 1993). Complacency has been linked to
359 higher reaction time (Bahner, Hüper, & Manzey, 2008; Manzey, Bahner, & Hüper,
360 2006), loss of situation awareness (Endsley & Kiris, 1995) and failures of detection
361 (Parasuraman et al., 1993). In our experiment, participants encountered no error during

362 the first three blocks. Given that the system never did any miss or error, participants may
363 have thought that it would remain perfectly reliable. In this context, their perception of
364 the required workload might evolve: since the automated system does not seem to require
365 their attention to function properly, participants would redirect their cognitive resources
366 towards more personal matters and mind-wander more. The higher perceived workload
367 under the manual condition supports our analysis. Moreover, this could explain why
368 participants, novice in supervising the system, exhibited an increase of MW frequency
369 only after some time, while pilots in Casner and Schooler (2015), who were pilots with
370 thousands of hours dealing with autopilot, mind wandered immediately without temporal
371 evolution. These evidences suggest a mediating influence of system familiarity in MW
372 frequency temporal evolution. This position would introduce a third possibility within the
373 overload/underload theory debate (Pattyn et al., 2008; Warm, Parasuraman, & Matthews,
374 2008). Although the task complexity does not change, the operator's perception could
375 evolve based on their trust in the system and their feelings toward the overall situation.
376 As pointed by Seli and colleagues (Seli, Carriere, & Smilek, 2015; Seli et al., 2016), there
377 is strong evidence that people can exert some control over their MW. This follows Casner
378 and Schooler's (2015) results, who demonstrated that cognitive resources freed by
379 automation in peaceful situations are not allocated to task planning, but rather to MW.
380 Our analysis is in line with studies that observed MW increase in a low probability signal
381 environment (Berthié et al., 2015; Casner & Schooler, 2015; Galera et al., 2012), with the
382 time elapsed performing the task (McVay & Kane, 2009; Smallwood, Baracaia, Lowe, &
383 Obonsawin, 2003; Smallwood, Riby, Heim, & Davies, 2006) and the view of
384 complacency as a multiple-task strategy (Bahner et al., 2008; Moray & Inagaki, 2000).

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

388 The second possible explanation is a decoupling of operators' attention from the
389 task at hand. When dealing with automation, operators give up their direct control over
390 the system for a monitoring role in the supervisory control loop (Moray, 1986; Sheridan,
391 1992). They may experience a loss of agency – i.e., the ability to feel in control (Wegner,
392 2002). Multiple studies pointed to a limit to the automation level beyond which users felt
393 less in control (Berberian, Sarrazin, Le Blaye, & Haggard, 2012; Coyle, Moore,
394 Kristensson, Fletcher, & Blackwell, 2012), leading to a form of disengagement from the
395 task at hand (Haggard, 2017). Interestingly, Szalma (2014) described a similar
396 disengagement when applying the Self-Determination Theory (Ryan & Deci, 2000) to
397 human-system interactions. The inability of a system to support autonomous behavior
398 may lower motivation and create an externalization of task goals – i.e. a process by which
399 operators rejects the value of a goal. In our experiment, since participants do not validate
400 but rather only acknowledge the system's actions, they could firstly experience a loss of
401 agency, their motivation would decrease, leading to a faint sense of responsibility. This
402 process chain could lead participants to reallocate cognitive resources from the task to
403 MW, unconsciously trying to optimize time and mental resources from their perspective.
404 Further studies are needed to distinguish the respective impacts of agency drop and
405 complacency on MW emergence.

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

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

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

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

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

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

469 In the near future, the massive use of automation in everyday systems will
470 reinforce the OOTL phenomenon. Our results show that automation increases MW
471 frequency after some time. The MW literature in ecological tasks already highlighted
472 how the phenomenon increases the risks in critical environments. Such results stress the
473 necessity to study in more detail the relation between MW and the OOTL performance
474 problem. Possible improvements include the study of reliability and complacency by
475 manipulating the number of conflicts and automation errors. Another possibility is to
476 highlight the impact of the operator's engagement in the task. Finally, perceived

477 workload is not to be overlooked. The use of electroencephalograms would allow
478 continuous measurement to precisely assess its impact over MW frequency. However,
479 such a protocol requires the influence of perceived workload and MW over neural
480 measures to be discriminated. Eventually, the expected outcome is to design automated
481 systems able to adapt themselves to operators' MW episodes. We hope that such a system
482 may enhance safety in critical automated environments.

483

484 **Acknowledgements**

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

488

489 **References**

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

- 500 4. Bastian, M., Lerique, S., Adam, V., Franklin, M. S., Schooler, J. W., & Sackur, J.
501 (2017). Language facilitates introspection: Verbal mind-wandering has privileged
502 access to consciousness. *Consciousness and Cognition*, *49*, 86–97.
503 <https://doi.org/10.1016/j.concog.2017.01.002>
- 504 5. Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-
505 Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
506 <https://doi.org/doi:10.18637/jss.v067.i01>
- 507 6. Baxter, G., Rooksby, J., Wang, Y., & Khajeh-Hosseini, A. (2012). The ironies of
508 automation ... still going strong at 30? In *Proceedings of ECCE 2012 Conference*.
509 Edinburgh, North Britain.
- 510 7. Berberian, B., Sarrazin, J.-C., Le Blaye, P., & Haggard, P. (2012). Automation
511 Technology and Sense of Control: A Window on Human Agency. *PLoS ONE*,
512 *7*(3), e34075. <https://doi.org/10.1371/journal.pone.0034075>
- 513 8. Berthié, G., Lemerrier, C., Paubel, P.-V., Cour, M., Fort, A., Galera, C., ...
514 Maury, B. (2015). The restless mind while driving: drivers' thoughts behind the
515 wheel. *Accident Analysis & Prevention*, *76*, 159–165.
516 <https://doi.org/10.1016/j.aap.2015.01.005>
- 517 9. Bixler, R., & D'Mello, S. (2014). Toward Fully Automated Person-Independent
518 Detection of Mind Wandering. In *User Modeling, Adaptation, and*
519 *Personalization* (pp. 37–48). Springer. Retrieved from
520 http://link.springer.com/chapter/10.1007/978-3-319-08786-3_4
- 521 10. Bixler, R., & D'Mello, S. (2015). Automatic gaze-based user-independent
522 detection of mind wandering during computerized reading. In *Proceedings of the*

- 523 *23rd User Modeling and User-Adapted Interactions conference* (pp. 31–43).
524 Dublin, Ireland. <https://doi.org/10.1007/s11257-015-9167-1>
- 525 11. Cabon, P., Coblenz, A., Mollard, R., & Fouillot, J. P. (1993). Human vigilance in
526 railway and long-haul flight operation. *Ergonomics*, *36*(9), 1019–1033.
527 <https://doi.org/10.1080/00140139308967974>
- 528 12. Casner, S. M., & Schooler, J. W. (2015). Vigilance impossible: Diligence,
529 distraction, and daydreaming all lead to failures in a practical monitoring task.
530 *Consciousness and Cognition*, *35*, 33–41.
531 <https://doi.org/10.1016/j.concog.2015.04.019>
- 532 13. Cegarra, J., & Morgado, N. (2009). Étude des propriétés de la version
533 francophone du NASATLX. In B. Cahour, F. Anceaux, & A. Giboin (Eds.), *Acte*
534 *du 5ème colloque de psychologie ergonomique (EPIQUE'2009)* (pp. 233–239).
535 Nice, France.
- 536 14. Cheyne, A. J., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy
537 of an error: A bidirectional state model of task engagement/disengagement and
538 attention-related errors. *Cognition*, *111*(1), 98–113.
539 <https://doi.org/10.1016/j.cognition.2008.12.009>
- 540 15. Coyle, D., Moore, J., Kristensson, P. O., Fletcher, P., & Blackwell, A. (2012). I
541 did that! Measuring users' experience of agency in their own actions. In
542 *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*
543 (pp. 2025–2034). Austin, TX: ACM. <https://doi.org/10.1145/1518701.1518766>
- 544 16. Cummings, M. L., Sasangohar, F., Thornburg, K. M., Xing, J., & D'Agostino, A.
545 (2010). Human-system interface complexity and opacity part i: literature review.

- 546 *Massachusetts Institute of Technology, Cambridge, MA*. Retrieved from
547 https://www.researchgate.net/profile/Farzan_Sasangohar/publication/263315585_
548 Human-
549 [system_interface_complexity_and_opacity/links/0deec53accc4413fb2000000.pdf](https://www.researchgate.net/profile/Farzan_Sasangohar/publication/263315585_system_interface_complexity_and_opacity/links/0deec53accc4413fb2000000.pdf)
- 550 17. Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. London,
551 England: Academic Press.
- 552 18. Dündar, C. (2015). *The effects of mind wandering on simulated driving*
553 *performance*. Retrieved from
554 <http://etd.lib.metu.edu.tr/upload/12618921/index.pdf>
- 555 19. Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The
556 Unengaged Mind: Defining Boredom in Terms of Attention. *Perspectives on*
557 *Psychological Science*, 7(5), 482–495.
558 <https://doi.org/10.1177/1745691612456044>
- 559 20. Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem
560 and level of control in automation. *Human Factors: The Journal of the Human*
561 *Factors and Ergonomics Society*, 37(2), 381–394.
- 562 21. Esterman, M., Noonan, S. K., Rosenberg, M., & DeGutis, J. (2013). In the Zone
563 or Zoning Out? Tracking Behavioral and Neural Fluctuations During Sustained
564 Attention. *Cerebral Cortex*, 23(11), 2712–2723.
565 <https://doi.org/10.1093/cercor/bhs261>
- 566 22. Faber, M., Bixler, R., & D’Mello, S. K. (2017). An automated behavioral measure
567 of mind wandering during computerized reading. *Behavior Research Methods*.
568 <https://doi.org/10.3758/s13428-017-0857-y>

- 569 23. Franklin, M. S., Broadway, J. M., Mrazek, M. D., Smallwood, J., & Schooler, J.
570 W. (2013). Window to the wandering mind: Pupillometry of spontaneous thought
571 while reading. *The Quarterly Journal of Experimental Psychology*, *66*(12), 2289–
572 2294. <https://doi.org/10.1080/17470218.2013.858170>
- 573 24. Fraulini, N. W., Hancock, G. M., Neigel, A. R., Claypoole, V. L., & Szalma, J. L.
574 (2017). A critical examination of the research and theoretical underpinnings
575 discussed in Thomson, Besner, and Smilek (2016). *Psychological Review*, *124*(4),
576 525–531. <https://doi.org/10.1037/rev0000066>
- 577 25. Fried, R., Schettlinger, K., & Borowski, M. (2014). robfilter: Robust Time Series
578 Filters (Version 4.1). Retrieved from [https://CRAN.R-](https://CRAN.R-project.org/package=robfilter)
579 [project.org/package=robfilter](https://CRAN.R-project.org/package=robfilter)
- 580 26. Galera, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribereau-
581 Gayon, R., ... Lagarde, E. (2012). Mind wandering and driving: responsibility
582 case-control study. *British Medical Journal*, *345*(dec13 8), e8105–e8105.
583 <https://doi.org/10.1136/bmj.e8105>
- 584 27. Gerbert, K., & Kemmler, R. (1986). The causes of causes: determinants and
585 background variables of human factor incidents and accidents. *Ergonomics*,
586 *29*(11), 1439–1453. <https://doi.org/10.1080/00140138608967257>
- 587 28. Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil
588 diameter tracks changes in control state predicted by the adaptive gain theory of
589 locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2),
590 252–269. <https://doi.org/10.3758/CABN.10.2.252>

- 591 29. Golchert, J., Smallwood, J., Jefferies, E., Seli, P., Huntenburg, J. M., Liem, F., ...
592 Margulies, D. S. (2016). Individual variation in intentionality in the mind-
593 wandering state is reflected in the integration of the default-mode, fronto-parietal,
594 and limbic networks. *NeuroImage*, *146*, 226–235.
595 <https://doi.org/10.1016/j.neuroimage.2016.11.025>
- 596 30. Gouraud, J., Delorme, A., & Berberian, B. (2017). Autopilot, Mind Wandering,
597 and the Out of the Loop Performance Problem. *Frontiers in Neuroscience*, *11*.
598 <https://doi.org/10.3389/fnins.2017.00541>
- 599 31. Grandchamp, R., Braboszcz, C., & Delorme, A. (2014). Oculometric variations
600 during mind wandering. *Frontiers in Psychology*, *5*.
601 <https://doi.org/10.3389/fpsyg.2014.00031>
- 602 32. Haggard, P. (2017). Sense of agency in the human brain. *Nature Reviews*
603 *Neuroscience*, *18*(4), 196–207. <https://doi.org/10.1038/nrn.2017.14>
- 604 33. Hancock, P. A. (2013). In search of vigilance: The problem of iatrogenically
605 created psychological phenomena. *American Psychologist*, *68*(2), 97–109.
606 <https://doi.org/10.1037/a0030214>
- 607 34. Hart, S. G., & Staveland, L. E. (1988). Development of NASA- TLX (Task Load
608 Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*,
609 139–203.
- 610 35. Head, J., & Helton, W. S. (2016). The troubling science of neurophenomenology.
611 *Experimental Brain Research*. <https://doi.org/10.1007/s00221-016-4623-7>
- 612 36. He, J., Becic, E., Lee, Y.-C., & McCarley, J. S. (2011). Mind Wandering Behind
613 the Wheel: Performance and Oculomotor Correlates. *Human Factors: The*

- 614 *Journal of the Human Factors and Ergonomics Society*, 53(1), 17–21.
615 <https://doi.org/10.1177/0018720810391530>
- 616 37. Hothorn, T., Bretz, F., Westfall, P., Herberger M., R., Schuetzenmeister, A., &
617 Schreibe, S. (2017). multComp: Simultaneous Inference in General Parametric
618 Models (Version 1.4-8). Retrieved from [https://CRAN.R-](https://CRAN.R-project.org/package=multcomp)
619 [project.org/package=multcomp](https://CRAN.R-project.org/package=multcomp)
- 620 38. Jeroski, J., Miller, M. E., Langhals, B., & Tripp, L. (2014). Impact of Vigilance
621 Decrement upon Physiology Measures. In *IIE Annual Conference. Proceedings*
622 (p. 1184). Institute of Industrial Engineers-Publisher. Retrieved from
623 <http://www.xcdsystem.com/iie2014/abstract/finalpapers/I471.pdf>
- 624 39. Kam, J. W. Y., Dao, E., Blinn, P., Krigolson, O. E., Boyd, L. A., & Handy, T. C.
625 (2012). Mind wandering and motor control: off-task thinking disrupts the online
626 adjustment of behavior. *Frontiers in Human Neuroscience*, 6.
627 <https://doi.org/10.3389/fnhum.2012.00329>
- 628 40. Kam, J. W. Y., Dao, E., Farley, J., Fitzpatrick, K., Smallwood, J., Schooler, J. W.,
629 & Handy, T. C. (2011). Slow fluctuations in attentional control of sensory cortex.
630 *Journal of Cognitive Neuroscience*, 23(2), 460–470.
- 631 41. Killingsworth, M. A., & Gilbert, D. T. (2010). A Wandering Mind Is an Unhappy
632 Mind. *Science*, 330(6006), 932–932. <https://doi.org/10.1126/science.1192439>
- 633 42. K. McIntire, L., P. McIntire, J., Mckinley, R., & Goodyear, C. (2014). Detection
634 of vigilance performance with pupillometry. In *Proceedings of the Symposium on*
635 *Eye Tracking Research and Applications* (pp. 167–174). Safety Harbor, FL, USA:
636 ACM.

- 637 43. Konishi, M., Brown, K., Battaglini, L., & Smallwood, J. (2017). When attention
638 wanders: Pupillometric signatures of fluctuations in external attention. *Cognition*,
639 *168*, 16–26. <https://doi.org/10.1016/j.cognition.2017.06.006>
- 640 44. Lawrence, M. A. (2016). ez: Easy Analysis and Visualization of Factorial
641 Experiments (Version 4.4-0). Retrieved from [https://CRAN.R-](https://CRAN.R-project.org/package=ez)
642 [project.org/package=ez](https://CRAN.R-project.org/package=ez)
- 643 45. Lemerrier, A. (2014). *Développement de la pupillométrie pour la mesure*
644 *objective des émotions dans le contexte de la consommation alimentaire*
645 (Psychologie). Université Paris 8.
- 646 46. Lerner, N., Baldwin, C., Higgins, J. S., Lee, J., & Schooler, J. (2015). Mind
647 Wandering While Driving: What Does it Mean and What do we do about it?
648 *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*,
649 *59*(1), 1686–1690. <https://doi.org/10.1177/1541931215591364>
- 650 47. Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual
651 search. *Quarterly Journal of Experimental Psychology*, *1*, 6–14.
652 <https://doi.org/10.1080/17470214808416738>
- 653 48. Mair, P., Schoenbrodt, F., & Wilcox, R. (2017). WRS2: Wilcox robust estimation
654 and testing.
- 655 49. Manzey, D., Bahner, J. E., & Hüper, A.-D. (2006). Misuse of automated aids in
656 process control: Complacency, automation bias and possible training
657 interventions. In *Proceedings of the Human Factors and Ergonomics Society*
658 *Annual Meeting* (Vol. 50, pp. 220–224). Los Angeles, CA: Sage Publications.

- 659 Retrieved from
660 <http://journals.sagepub.com/doi/abs/10.1177/154193120605000303>
- 661 50. McMillan, R. L., Kaufman, S. B., & Singer, J. L. (2013). Ode to positive
662 constructive daydreaming. *Frontiers in Psychology, 4*.
663 <https://doi.org/10.3389/fpsyg.2013.00626>
- 664 51. McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: working
665 memory capacity, goal neglect, and mind wandering in an executive-control task.
666 *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(1),
667 196.
- 668 52. Methot, L. L., & Huitema, B. E. (1998). Effects of signal probability on
669 individual differences in vigilance. *Human Factors: The Journal of the Human*
670 *Factors and Ergonomics Society, 40*(1), 102–110.
- 671 53. Mittner, M., Boekel, W., Tucker, A. M., Turner, B. M., Heathcote, A., &
672 Forstmann, B. U. (2014). When the Brain Takes a Break: A Model-Based
673 Analysis of Mind Wandering. *Journal of Neuroscience, 34*(49), 16286–16295.
674 <https://doi.org/10.1523/JNEUROSCI.2062-14.2014>
- 675 54. Moray, N. (1986). Monitoring behavior and supervisory control. In K. Boff (Ed.),
676 *Handbook of perception and human performance* (Vol. 2, pp. 1–51). New York,
677 NY: Wiley.
- 678 55. Moray, N., & Inagaki, T. (2000). Attention and complacency. *Theoretical Issues*
679 *in Ergonomics Science, 1*(4), 354–365.
680 <https://doi.org/10.1080/14639220052399159>

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

- 704 63. Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye Movements
705 During Mindless Reading. *Psychological Science*, *21*(9), 1300–1310.
706 <https://doi.org/10.1177/0956797610378686>
- 707 64. Revelle, W. (2017). psych: Procedures for Personality and Psychological
708 Research (Version 1.7.5). Evanston, Illinois, USA: Northwestern University.
709 Retrieved from <https://CRAN.R-project.org/package=psych>
- 710 65. Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997).
711 “Oops!”: Performance correlates of everyday attentional failures in traumatic
712 brain injured and normal subjects. *Neuropsychologia*, *35*(6), 747–758.
- 713 66. RStudio Team. (2015). RStudio: Integrated Development for R (Version 1.0.143).
714 Boston, MA, USA: RStudio Inc. Retrieved from <http://www.rstudio.com/>
- 715 67. Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation
716 of intrinsic motivation, social development, and well-being. *American*
717 *Psychologist*, *55*(1), 68.
- 718 68. Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises.
719 *Handbook of Human Factors and Ergonomics*, *2*, 1926–1943.
- 720 69. Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., &
721 Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering
722 mind. *Trends in Cognitive Sciences*, *5*, 319–326.
723 <https://doi.org/10.1016/j.tics.2011.05.006>
- 724 70. Seli, P., Carriere, J. S. A., & Smilek, D. (2015). Not all mind wandering is created
725 equal: dissociating deliberate from spontaneous mind wandering. *Psychological*
726 *Research*, *79*(5), 750–758. <https://doi.org/10.1007/s00426-014-0617-x>

- 727 71. Seli, P., Risko, E. F., & Smilek, D. (2016). On the necessity of distinguishing
728 between unintentional and intentional mind wandering. *Psychological Science*,
729 27(5), 685–691. <https://doi.org/10.1177/0956797616634068>
- 730 72. Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*.
731 Cambridge: MIT Press.
- 732 73. Smallwood, J. (2013). Distinguishing how from why the mind wanders: A
733 process–occurrence framework for self-generated mental activity. *Psychological*
734 *Bulletin*, 139(3), 519–535. <https://doi.org/10.1037/a0030010>
- 735 74. Smallwood, J., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task
736 unrelated thought whilst encoding information. *Consciousness and Cognition*,
737 12(3), 452–484. [https://doi.org/10.1016/S1053-8100\(03\)00018-7](https://doi.org/10.1016/S1053-8100(03)00018-7)
- 738 75. Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek,
739 M. D., ... Schooler, J. W. (2011). Pupillometric Evidence for the Decoupling of
740 Attention from Perceptual Input during Offline Thought. *PLoS ONE*, 6(3),
741 e18298. <https://doi.org/10.1371/journal.pone.0018298>
- 742 76. Smallwood, J., Riby, L., Heim, D., & Davies, J. B. (2006). Encoding during the
743 attentional lapse: Accuracy of encoding during the semantic sustained attention to
744 response task. *Consciousness and Cognition*, 15(1), 218–231.
745 <https://doi.org/10.1016/j.concog.2005.03.003>
- 746 77. Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological*
747 *Bulletin*, 132(6), 946–958. <https://doi.org/10.1037/0033-2909.132.6.946>
- 748 78. Smallwood, J., & Schooler, J. W. (2015). The Science of Mind Wandering:
749 Empirically Navigating the Stream of Consciousness. *Annual Review of*

- 750 *Psychology*, 66(1), 487–518. <https://doi.org/10.1146/annurev-psych-010814->
751 015331
- 752 79. Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010a). Failures of sustained
753 attention in life, lab, and brain: Ecological validity of the SART.
754 *Neuropsychologia*, 48(9), 2564–2570.
755 <https://doi.org/10.1016/j.neuropsychologia.2010.05.002>
- 756 80. Smilek, D., Carriere, J. S. A., & Cheyne, J. A. (2010b). Out of Mind, Out of
757 Sight: Eye Blinking as Indicator and Embodiment of Mind Wandering.
758 *Psychological Science*, 21(6), 786–789.
759 <https://doi.org/10.1177/0956797610368063>
- 760 81. Szalma, J. L. (2014). On the Application of Motivation Theory to Human
761 Factors/Ergonomics: Motivational Design Principles for Human–Technology
762 Interaction. *Human Factors*, 56(8), pp. 1453–1471.
763 <https://doi.org/10.1177/0018720814553471>
- 764 82. Szalma, J. L., & Matthews, G. (2015). Motivation and Emotion in Sustained
765 Attention.
- 766 83. Szalma, J. L., Warm, J. S., Matthews, G., Dember, W. N., Weiler, E. M., Meier,
767 A., & Eggemeier, F. T. (2004). Effects of sensory modality and task duration on
768 performance, workload, and stress in sustained attention. *Human Factors: The*
769 *Journal of the Human Factors and Ergonomics Society*, 46(2), 219–233.
- 770 84. Thomson, D. R., Besner, D., & Smilek, D. (2016). A critical examination of the
771 evidence for sensitivity loss in modern vigilance tasks. *Psychological Review*,
772 123(1), 70–83. <https://doi.org/10.1037/rev0000021>

- 773 85. Thomson, D. R., Seli, P., Besner, D., & Smilek, D. (2014). On the link between
774 mind wandering and task performance over time. *Consciousness and Cognition*,
775 27, 14–26. <https://doi.org/10.1016/j.concog.2014.04.001>
- 776 86. Uzzaman, S., & Joordens, S. (2011). The eyes know what you are thinking: Eye
777 movements as an objective measure of mind wandering. *Consciousness and*
778 *Cognition*, 20(4), 1882–1886. <https://doi.org/10.1016/j.concog.2011.09.010>
- 779 87. Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance Requires Hard
780 Mental Work and Is Stressful. *Human Factors: The Journal of the Human*
781 *Factors and Ergonomics Society*, 50(3), 433–441.
782 <https://doi.org/10.1518/001872008X312152>
- 783 88. Wegner, D. M. (2002). *The Illusion of Conscious Will* (MA: MIT Press).
784 Cambridge.
- 785 89. Wickham, H. (2007). Reshaping data with the reshape package. *Journal of*
786 *Statistical Software*, 21(12).
- 787 90. Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-
788 *Verlag New York*.
- 789 91. Wiener, E. L. (1988). *Cockpit automation*. National Aeronautics and Space
790 Administration. Retrieved from <http://ntrs.nasa.gov/search.jsp?R=19890047073>
- 791 92. Wise, J. A., Tilden, D. S., Abbott, D., Dyck, J., & Guide, P. (1994). Managing
792 automation in the cockpit. In *Proceedings of the 24th International Conference of*
793 *the International Federation of Airworthiness*. Lisbon, Portugal. Retrieved from
794 http://www.faa.gov/training_testing/training/media/cfit/volume2/pdf/pages/page5
795 [_07.pdf](http://www.faa.gov/training_testing/training/media/cfit/volume2/pdf/pages/page5_07.pdf)

- 796 93. Yanko, M. R., & Spalek, T. M. (2014). Driving With the Wandering Mind: The
797 Effect That Mind-Wandering Has on Driving Performance. *Human Factors: The*
798 *Journal of the Human Factors and Ergonomics Society*, 56(2), 260–269.
799 <https://doi.org/10.1177/0018720813495280>
800

Figure 1

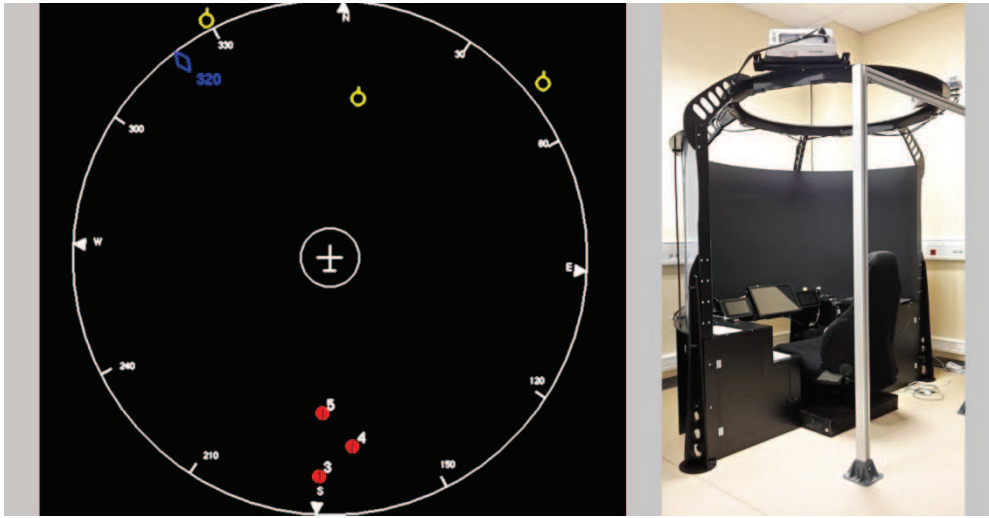


Figure 1. Screenshot of the LIPS interface and the environment. One of the screen is used for the task and the other one for questionnaire probes. For the task, the plane in the center is static and the surround (yellow and red numbered symbols) are moving. During, left and right avoidance maneuvers,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

Figure 2

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

Tâche
Quelque chose lié à la tâche
Quelque chose non-lié à la tâche
Une distraction externe

[Sauvegarder et quitter](#)

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

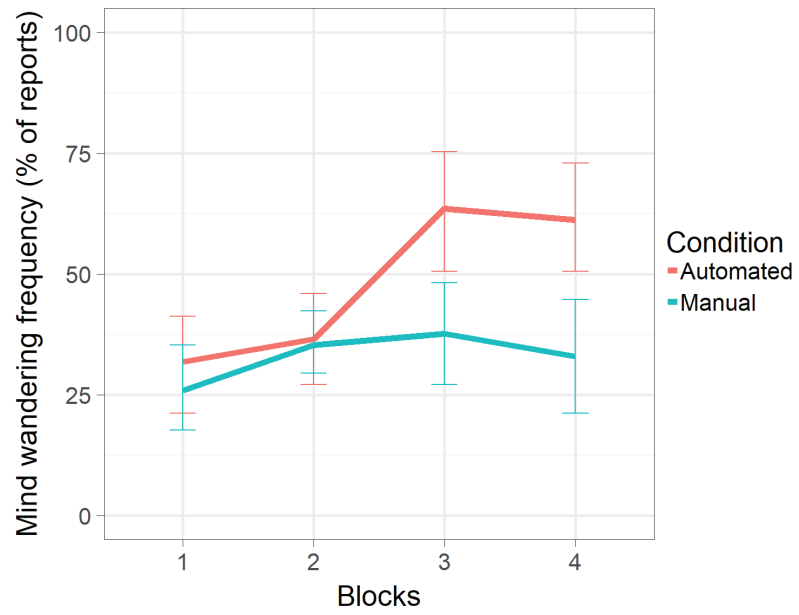


Figure 3: MW frequency evolution according to the condition. Error bars show the 95% confidence intervals based on bootstrap

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	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	$\chi^2(1)$	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)	$2.87 \cdot 10^{-1}$	$6.65 \cdot 10^{-1}$	$3.22 \cdot 10^{-1}$	$6.52 \cdot 10^{-1}$	0.08	.774
Blink frequency (blink/s)	$6.90 \cdot 10^{-2}$	$1.10 \cdot 10^{-1}$	$5.43 \cdot 10^{-2}$	$9.81 \cdot 10^{-2}$	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