A Preliminary Study on the Influence of Automation over Mind Wandering Frequency in Sustained Attention

Jonas Gouraud
Information Processing and Systems Department
ONERA, Base Aérienne 701
13661, Salon Air, France
contact@jonasgouraud.com

Arnaud Delorme
Swartz Center for Computational Neurosciences, University of California
San Diego, La Jolla, CA93092-0559
Centre de Recherche Cerveau et Cognition - UMR 5549
Pavillon Baudot, Hopital Purpan, BP 25202
31052 Toulouse Cedex 3, France
arno@cnrs.fr

Bruno Berberian
Information Processing and Systems Department
ONERA, Base Aérienne 701
13661, Salon Air, France
+33490175588
bruno.berberian@onera.fr

ABSTRACT

To satisfy the increasing demand for safer critical systems, engineers have integrated higher levels of automation. In the context of airplane autopilot, time saved by automation, which should normally be used to plan the flight, might instead be filled by task-unrelated thoughts, or mind wandering (MW). We observed the impact of automation on MW in an operational environment. Participants were required to either avoid incoming obstacles by controlling the movements of an aircraft on a 2D radar screen or monitor an automated system performing the same task. Participants' propensity to mind wander increased with the time spent doing the task. Moreover, the time spent MW increased with automation in a significant manner. The NASA TLX, a measure of perceived workload, highlighted the influence of automation over perceived workload. Moreover, TLX scores were not correlated with MW propensity. This study shows a significant influence of automation over MW, which was not due to workload effects or task interactions.

CCS CONCEPTS

• **User/Machine Systems** → Human factors • **Organizational Impacts** → Automation

KEYWORDS

ACM proceedings, aeronautics, mind wandering, automation, perceived workload, complacency

ECCE 2017, September 19-22, 2017, Umeå, Sweden

ACM ISBN 978-1-4503-5256-7/17/09$15.00
https://doi.org/10.1145/3121283.3121306

ACM Reference format:

G. Gubbiotti, P. Malagò, S. Fin, S. Tacchi, L. Giovannini, D. Bisero, M. Madami, and G. Carlotti. 1997. SIG Proceedings Paper in word Format. In *Proceedings of ACM Woodstock conference, El Paso, Texas USA, July 1997 (WOODSTOCK'97),* 4 pages.
DOI: 10.1145/123 4

1 INTRODUCTION

In order to continuously improve system safety, critical systems industry makes an extensive use of automation. In cockpits [14], in cars [8], automation has been introduced to increase performance and respond to new safety requirements. Unfortunately, while implementing higher levels of automation improves the efficiency and capacity of a system, it also creates new challenges for human operators.

Researchers have shown that vigilance failure is a key component in many accidents where automation is involved [1]. Reports in aviation security have notably illustrated the role of human error when interacting with highly automated systems. Gerbert and Kemmler [4] studied 1448 German aviators' anonymous responses to questionnaires about automation-related incidents and reported that failure of vigilance was the largest contributor to human error. Indeed, several studies show that efficient sustained attention over hours cannot be achieved [7]. If research on vigilance suggests that time on task decreases significantly our ability to discriminate infrequent signals [12], vigilance failure also encompasses another reality when dealing with automation: the complacency experienced by operators dealing with highly reliable automated systems [9]. Complacency is created by blind reliance on a system leading the operator to think that it is more competent than it actually is. As operators have the feeling that the system does not require them to work efficiently, they instinctively lower cognitive resources allocated to monitoring it.

Such a context favors the occurrence of mind wandering (MW) episodes. MW is "the human mind propensity to drift away from the task at hand towards unrelated inner thoughts, fantasies, and feelings" [11]. It often occurs without awareness on the operator's part. As it diverts operator's attention from his primary task, it could play an important role in vigilance failure observed in highly reliable automated environments. Casner and Schooler [2] studied MW evolution in the cockpit. Their results did not show a significant correlation between the level of automation and the frequency of MW reports. However MW frequency increased when the pilot did not interact with his environment. This could be due to a decrease in vigilance as well as a difference in complexity as compared when they are directly interacting with the system, and might not necessarily be linked to automation. Moreover, they demonstrated that in a situation where everything seems under control, participants mind wandered more. Whereas monitoring tasks require high levels of mental resources, supervising ultra-reliable systems encourages 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 monitoring, was instead filled by task-unrelated thoughts. In other words, when experiencing complacency, operators' cognitive resources not allocated to monitoring might be redirected towards MW. Such assertion is supported by observed MW increase in low probability signal environment [2,7,14] and with time on task [6]. As MW happens regularly – if not always – without intention and awareness, it could further impair the ability to respond to rare critical events.

We believe automation might increase the number of MW reports in highly automated system as compared to manual operation of a system. Our experiment addresses this hypothesis.

2 METHOD

2.1 Participants

6 participants (all male) performed the experiment (age ranging from 24 to 42 years-old; M = 28.2, SD = 6.9). The participants enrolled in this study were volunteers from the ONERA organization. All participants had normal or corrected-to-normal visual acuity.

2.2 Task

We used the LIPS environment developed at the ONERA organization to program our experiment (see fig. 1). An unmanned air vehicle (UAV) stayed at the center of a 2D radar 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 and 5 obstacles, including one on the trajectory. The participants were instructed to control the movements of the UAV to avoid obstacles.

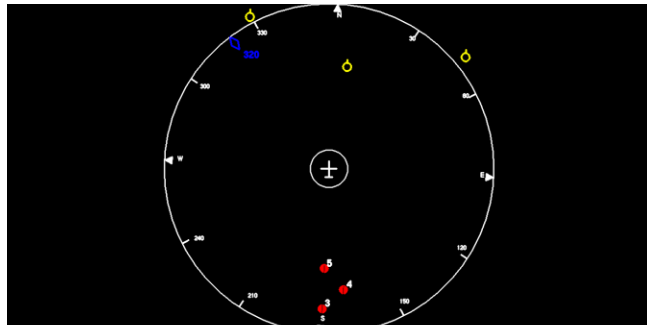


Figure 1. Screenshot of the LIPS in automated mode

Two conditions were proposed. The first one was the "manual" condition and required participants to manually avoid obstacles. The system detected the obstacle 13s before impact. Once collision was detected, an orange circle appeared around the UAV and the participant could initiate an avoidance spelling. Participants were able to choose the way in which they wished to avoid the obstacle by pressing "Evitement Gauche" (left maneuver) or "Evitement Droite" (right maneuver). Depending on time to impact, the optimal UAV trajectory was followed. Each obstacle had a safe circle similar to the one of the UAV (see fig. 1). A proximity warning – i.e. orange circle around both UAVs and obstacle with a message "Collision" – was displayed if the UAV penetrated inside. Any trial with a proximity warning was marked as failed. To resume the initial trajectory, the participant had to press the "Retour trajectoire" (return to original trajectory) button. If no action was taken 16 seconds after the first change in trajectory, the aircraft resumed automatically 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 press the "Acquittement" (acknowledge) button to acknowledge automation decisions as soon as they saw it. It had to be done each time the system changed the trajectory – twice per trial. A feedback message was displayed to the participants. The acknowledgement ensured participants would have the same number of interactions with the system as in manual mode. Finally, if participants detected a conflict that would result in a safety threat – obstacle too close to the UAV or even collision – they were instructed to press 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.

2.3 Procedure

Participants were explicitly instructed that detection accuracy was overall more important than speed on button presses. Each participant performed the two conditions on two separate days. For the subjects 1, 3 and 5, the first day started with an explanation of the task, followed by a 10-minutes training and a 45-minutes session for the first condition. The second day consisted of a 10-minutes training and a 45-minutes session in the second condition. Conditions were counterbalanced for subjects 2, 4 and 6. Each session contained 60 clusters of obstacles, each cluster randomly including between one and five obstacles. 21 questionnaires were answered in each condition (see below). In the automated condition, participants encountered one system error (where they had to press "Changement d'altitude" button) in the training session, and another during the actual session (at the end of the third block). Participants were informed of the duration of the experiment and could see the computer clock when the questionnaire was not displayed.

2.4 Online Questionnaire

Every 2 minutes, on average, a questionnaire appeared on a secondary screen next to the main task. Participants were asked to fill it as soon as it appeared. The experiment was not paused, however participants were informed that any error or miss during this interval would not be taken into account for the overall performances. We also stressed that the questionnaire was only for information purpose and was not part of the evaluation. This ensured that participants would not be reluctant to report inattention. The questionnaire had the following questions (originally in French, translated here to English): "When this questionnaire appeared, where was your attention directed?" Answers could be "On task" (e.g. thinking about the next obstacle, the decision to make), "Something related to the task (e.g. thinking about performance, interface items, last trial), "Something unrelated to the task" (e.g. thinking about a random memory, a planning, a body sensation) or "External distraction" (e.g. conversation, noise). The preceding examples were given to participants to illustrate each category.

2.5 Post-task Questionnaire

We used a validated French version of the NASA TLX [3] questionnaire to evaluate workload along several dimensions. This questionnaire includes questions pertaining to mental workload, time pressure, physical strain, effort, frustration, and performance satisfaction. Participants were asked to answer each question using a horizontal line, ranging from "low" to "high" on a scale from 0 to 20. The NASA TLX was only filled at the end of each session to limit disruption during the task itself. A TLX questionnaire proposed at each block would allow a more precise workload monitoring. However MW would have been decreased artificially by this disruption. Similarly, online measures of workload may be difficult to use because of the necessity to distinguish physiological influence of MW and workload.

3 RESULTS AND DISCUSSION

3.1 Mind Wandering Frequency

We split the 45 sessions into 4 blocks containing 5 reports each. MW propensity was calculated in percentage of all reports in the block. Participants reported on average 2.65 MW episodes (SD = 1.28) per block, corresponding to 53% of time on task spent in MW. This rate is consistent with what is being reported in the literature [5].

All participants exhibited a general increase in their MW frequency (see fig. 2) as the task progressed, consistent with other studies [5]. As the environment in both conditions is monotonous and actions are seldom required, participants have plenty of time to mind wander. However, we see in fig. 3

Figure 3: MW frequency evolution for each condition

that this evolution differs according to the level of automation. In manual mode, participants' increase is approximately linear. By contrast, MW frequency in automated mode exhibits a steeper slope between the second and the third blocks than between any other blocks.

Given the limited amount of data, we performed a Friedman's ANOVA on block number for automated and manual conditions separately. In the automated condition, there was a significant difference between each block, χ^2 (3) = 13.98, p < .001. This result is supported by the concordance between subjects, C = .777. In the manual condition, no significant difference between blocks was observed, χ^2 (3) = 4.36, p = .224, a result supported by the low concordance, C = .242. Post-hoc tests were conducted using the Wilcoxon's signed-rank test. We explored the difference between conditions by taking together block 1 and 2 on one hand, and block 3 and 4 on the other hand. The first 2 blocks did not exhibit a significant difference between automated (Mdn = 0.40) compared to the manual condition (Mdn = 0.40), p = .246, r = -.167. However, the last two blocks did show a significantly higher MW frequency in the automated (Mdn = 0.80) compared to the manual condition (Mdn = 0.40), p = .002, r = -.444.

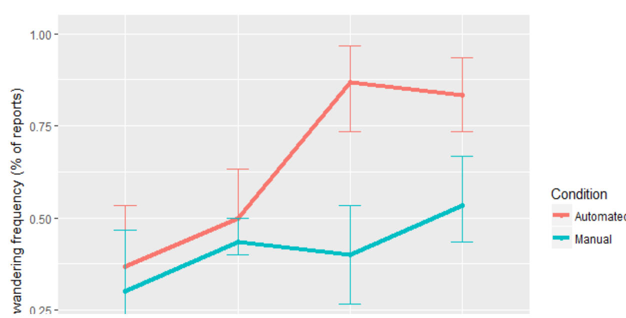
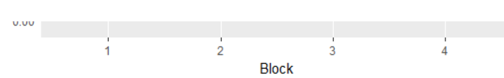


Figure 2. MW frequency evolution for each subject



Preliminary results show that automation increases MW propensity compared to manual conditions, with an influence noticeable only after half an hour on task. This effect could be due to complacency. During the first three blocks, participants encounter no conflict requiring an action, therefore their only task is to acknowledge actions of the automation. The first and second block may only show a normal MW increase, while the participants learn the reliability of automation. However, during the third block, complacency may emerge and lower cognitive resources allocated to the task. These resources may then be used to think about unrelated matters, thus increasing MW reports frequency. On the fourth block, as a system error is encountered in the automated condition, complacency is lower and MW frequency slightly decreases. This is in line with Casner and Schooler [2], who demonstrated that cognitive resources freed by automation are not allocated to flight planning but rather to MW.

3.2 NASA TLX Scores

Fig. 4 shows NASA TLX questionnaire scores (M = 5.31, SD = 3.74) for each subject and each session. Shapiro-Wilk's test indicated that the assumption of normality had been violated for the TLX values, W = 0.929, p < .001. The Wilcoxon signed-rank values in automated (Mdn = 3.00) and manual (Mdn = 7.00) modes, p = .004, r = -.338.

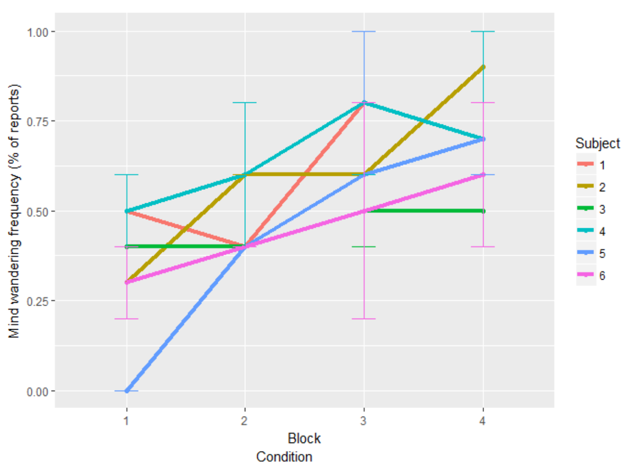


Figure 4: NASA TLX Scores for each subject

A tendency exists for the automation condition to be perceived as necessitating a lower workload. As our system only proposes one situation in each automated session when a human intervention is required to correct a system error, this drop in perceived workload could be explained by complacency. When dealing with ultra-safe systems, operators can overtrust the system and let it handle the situation without proper monitoring. As less cognitive resources are dedicated to the task, workload is perceived as lower. However, both conditions produced low TLX scores. This is somewhat contradictory with the literature presenting monitoring as a stressful and demanding task [13].

4 CONCLUSIONS

Preliminary results show an increase of MW correlated with time on task. This relation is also influenced by automation, which produces a higher propensity of MW than in the manual condition. Simultaneously, automation lowered perceived workload at the end of the sessions. Collecting data on more subjects will allow us to rule on this hypothesis. We will also investigate the parameters influencing the time needed for an operator to experience complacency.

Possible future directions include using TLX questionnaire after each block to measure perceived workload and see if a significant difference is observed and correlated with MW propensity. Another possibility is to investigate MW psychophysiological markers to identify possible markers of complacency. Oculometry already demonstrated promising results in this regard [10].

ACKNOWLEDGMENTS

We thank the Direction Générale de l'Armement (DGA) for their financial support to the first author. This work has been supported by a grant from ANR (Young researcher program – ANR-15-CE26-0010-01).

REFERENCES

1. René Amalberti. 1999. Automation in aviation: A human factors perspective. *Handbook of aviation human factors*: 173–192.
2. Stephen M. Casner and Jonathan W. Schooler. 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
3. Julien Cegarra and Nicolas Morgado. 2009. Étude des propriétés de la version francophone du NASATLX. In *Communication présentée à la cinquième édition du colloque de psychologie ergonomique (Epique)*.
4. Karl Gerbert and Reiner Kemmler. 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
5. Julia WY Kam, Elizabeth Dao, James Farley, Kevin Fitzpatrick, Jonathan Smallwood, Jonathan W. Schooler, and Todd C. Handy. 2011. Slow fluctuations in attentional control of sensory cortex. *Journal of cognitive neuroscience* 23, 2: 460–470.
6. Jennifer C. McVay and Michael J. Kane. 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.
7. Laura L. Methot and Bradley E. Huitema. 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.
8. Frederik Naujoks, Christian Purucker, and Alexandra Neukum. 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
9. Raja Parasuraman, Robert Molloy, and Indramani L. Singh. 1993. Performance Consequences of Automation-Induced "Complacency." *The International Journal of Aviation Psychology*.
10. Jonathan Smallwood, Kevin S. Brown, Christine Tipper, Barry Giesbrecht, Michael S. Franklin, Michael D. Mrazek, Jean M. Carlson, and Jonathan W. Schooler. 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

11. Jonathan Smallwood and Jonathan W. Schooler. 2006. The restless mind. *Psychological Bulletin* 132, 6: 946–958. https://doi.org/10.1037/0033-2909.132.6.946
12. Joel S. Warm. 1984. *Sustained attention in human performance*.
13. Joel S. Warm, Raja Parasuraman, and Gerald Matthews. 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
14. John A. Wise, Donald S. Tilden, David Abbott, Jennifer Dyck, and Patrick Guide. 1994. Managing automation in the cockpit. In *International federation of airworthiness, 24th international conference*. Retrieved October 29, 2015 from http://www.faa.gov/training_testing/training/media/cfit/volume2/pdf/pages/page5_07.pdf