



HAL
open science

Supervision de système automatisé : effet du type d'automatisation et du niveau de fiabilité sur les performances humaines.

Eugénie Avril, Julien Cegarra, Liên Wioland, Jordan Navarro

► **To cite this version:**

Eugénie Avril, Julien Cegarra, Liên Wioland, Jordan Navarro. Supervision de système automatisé : effet du type d'automatisation et du niveau de fiabilité sur les performances humaines.. EPIQUE 2021, Jul 2021, Lille, France. <hal-03274678>

HAL Id: hal-03274678

<https://hal.science/hal-03274678v1>

Submitted on 30 Jun 2021

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.



HAL Authorization

Supervision de système automatisé : effets du type d'automatisation et du niveau de fiabilité sur les performances humaines

Eugénie Avril

Laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et Sociales - EA 7278
Université Côte d'Azur, 06357 Nice (France)
eugenie.avril@univ-cotedazur.fr

Julien Cegarra

Laboratoire Science de la Cognition, Technologie, Ergonomie (SCoTE) – EA 7420
Université de Toulouse, INU Champollion, 81000 Albi (France)
julien.cegarra@univ-jfc.fr

Liên Wioland

INRS, Département Homme au Travail
INRS, 54519 Vandœuvre-lès-Nancy (France)
lien.wioland@inrs.fr

Jordan Navarro

Laboratoire d'Etudes des Mécanismes Cognitifs – EA 3082
Université des Lumières Lyon 2, 69676 Bron (France)
jordan.navarro@univ-lyon2.fr

Catégorie de soumission : communication orale

RÉSUMÉ

Cette étude a pour but d'étudier la supervision de système lors de l'utilisation de deux types de systèmes automatisés : statique ou adaptatif. Notre objectif est d'étudier ces derniers dans des situations où la fiabilité n'est pas parfaite. Nous avons confronté 80 participants à trois tâches du *Multi Attribute Task Battery* (MATB). Nous avons manipulé quatre niveaux de fiabilité différentes et deux types de systèmes automatisés. Un système adaptatif avec des sessions de contrôle manuel planifiées, ainsi qu'un système statique, où le contrôle ne changeait jamais. Nos résultats ont montré que les participants ne mettaient pas en place les mêmes stratégies d'allocations visuelles d'attention selon les niveaux de fiabilité du système. Sous un système statique, les participants passaient plus de temps sur la tâche avec le niveau 0% de fiabilité, que dans le niveau 100% où la surveillance était réduite. Sous un système adaptatif, le comportement visuel était le même mais moins prononcé. Nos données ont également montré des résultats intéressants dans les sessions manuelles où les différents niveaux de fiabilité n'entraînent pas le même comportement visuel des participants lorsqu'ils retournent en supervision de système. Le niveau de fiabilité 56.25% a entraîné plus d'implication visuelle des participants durant les sessions automatisées, suivant une session de contrôle manuelle.

MOTS-CLÉS

Fiabilité, Interaction homme-machine, Allocation de tâches, Micromonde, Eye-tracker

1 INTRODUCTION

Avec l'évolution croissante des systèmes automatisés dans notre quotidien, les opérateurs humains sont amenés à superviser des systèmes accomplissant des tâches plutôt qu'effectuer celles-ci par eux-mêmes (Sheridan, 1970, 1987). Cette nouvelle répartition des tâches a déclenché de nouvelles conséquences d'interaction homme-machine telle que l'insuffisance de surveillance (Parasuraman & Manzey, 2010). Ce phénomène est souvent considéré comme un contentement des opérateurs dans des situations d'interaction avec des systèmes hautement fiables (Bagheri & Jamieson, 2004). Il est défini comme une faible suspicion envers le système automatisé menant à une surveillance sous-optimale de celui-ci (Wiener, 1981 ; Moray & Inagaki, 2000). Les opérateurs humains ont besoin d'un minimum de stimulation durant la supervision pour rester dans la boucle de contrôle (Endsley & Kiris, 1995). D'ailleurs, ceux-ci ont tendance à privilégier l'accomplissement manuel des tâches plutôt que l'automatisation afin de réduire leur ennui (Navarro & Osiurak, 2015).

Afin de maintenir les opérateurs humains dans la boucle de décision, différents types de systèmes automatisés ont été pensés par les concepteurs. Le système dit « adaptatif » permet de redonner le contrôle temporairement à l'opérateur humain (Kaber & Endsley, 2004). Il peut être invoqué par le système (adaptatif) ou par l'opérateur humain lui-même (adaptable). Parasuraman, Mouloua, et Singh (1996) ont étudié deux types de systèmes adaptatifs en utilisant un micromonde du pilotage aéronautique, le *Multi Attribute Task Battery* (Santiago-Espada et al., 2011). Ces systèmes alternaient entre des sessions automatisées et des sessions manuelles. Les participants ont commencé par une session de 10 minutes au cours de laquelle toutes les tâches MATB étaient effectuées manuellement. Après cette session, ils ont effectué trois sessions de 30 minutes au cours desquelles la tâche de surveillance a été automatisée (à 56.25% de fiabilité) tandis que les deux autres tâches sont restées sous contrôle manuel. Dans le premier système adaptatif dit "basé sur le modèle", le contrôle revenait toujours au participant (session manuelle) après 40 minutes sous automatisation. Le deuxième système adaptatif était basé sur la performance. Les participants recevaient un contrôle manuel aux moments prévus si la performance était inférieure à un taux spécifique (55%). Les résultats ont montré une augmentation similaire des performances (c'est-à-dire du taux de détection des dysfonctionnements) pour les deux systèmes automatisés adaptatifs.

Bien que la mise en place de sessions manuelles réduit l'insuffisance de surveillance, il y a peu d'études impliquant des systèmes automatisés adaptatifs en situation de problèmes de fiabilité (Oakley, Mouloua, & Hancock, 2003).

L'objectif principal de cette étude est d'étudier l'impact de la fiabilité de système adaptatif sur les performances humaines, par rapport à un système statique, en cas de dysfonctionnements.

Nos hypothèses étaient (1) qu'une augmentation du niveau de fiabilité conduirait à une insuffisance de surveillance dans la condition de système statique, se traduisant par une diminution du temps passé à examiner la tâche automatisée (Parasuraman & Manzey, 2010); (2) Cette inefficacité de surveillance devrait être atténuée dans les conditions impliquant un système adaptatif (Parasuraman, Mouloua, & Singh, 1996).

METHODE

1.1 Participants

80 participants de l'Institut National Universitaire Jean François Champollion (Albi) ont été sollicités pour réaliser cette expérience. Les participants étaient des étudiants âgés de 18 à 28 ans (62 femmes et 18 hommes) avec une moyenne d'âge de 19.9. Ils avaient tous une vision correcte.

1.2 Matériel et Protocole

1.2.1 The Multi Attribute Task Battery

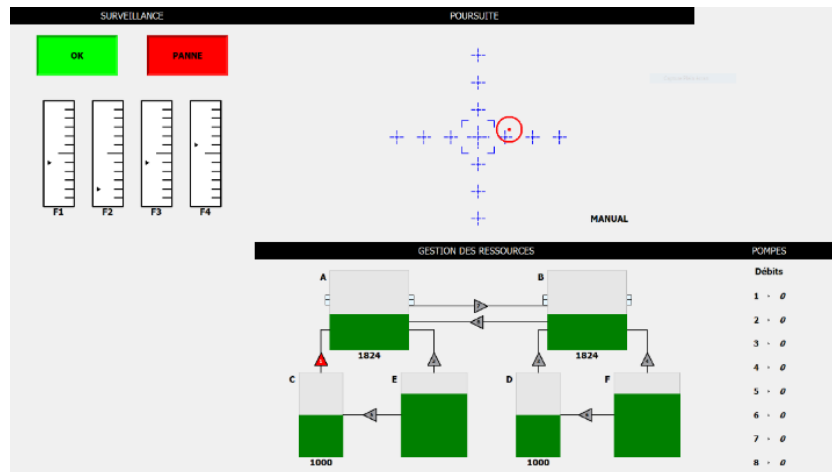


Figure 1. Capture d'écran du OpenMATB

Le MATB est un micro-monde composé de plusieurs tâches présentées simultanément. Une version modifiée opensource du MATB, intitulée *OpenMATB* a été conçue afin d'intégrer la synchronisation avec l'*eye-tracker* (Cegarra, Valéry, Avril, Calmettes, & Navarro, 2020). La tâche de cible est une tâche qui consistait à maintenir un curseur dans une zone cible en utilisant un Joystick. La tâche de gestion de ressources est une tâche de raisonnement où le participant doit maintenir des niveaux cibles de liquide dans des réservoirs qui se vident continuellement. La tâche de surveillance est composée de 4 jauges dans lesquelles un curseur fluctue. Il peut arriver que le curseur se bloque parfois et celui-ci doit être débloqué par le participant en appuyant sur des touches spécifiques du clavier. Dans notre étude, la tâche de surveillance était automatisée. Le curseur fluctuait au centre de chaque jauge. Lorsque le curseur se bloquait dans la partie supérieure ou inférieure d'une des jauges, le voyant « panne » s'allumait en rouge pour indiquer un problème et le système automatisé débloquent le curseur automatiquement. Lorsqu'il y avait une défaillance de ce système automatisé, le curseur ne se débloquent pas et le voyant rouge ne s'allumait pas. Le participant devait appuyer sur la touche F1 du clavier si le curseur était bloqué dans la première jauge ; sur F2 (jauge 2) ; sur F3 (jauge 3) ; F4 (jauge 4). Le curseur se débloquent automatiquement au bout de 10 secondes lorsque le participant n'agissait pas.

1.2.2 Plan expérimental

Les variables indépendantes étaient composées du type de systèmes automatisés (deux modalités : adaptatif ou statique) et du niveau de fiabilité de ces systèmes (quatre modalités : 0% ; 56.25% ; 87.5 % ; 100 % de fiabilité). Un participant était confronté à un seul type de système automatisé et à un seul niveau de fiabilité. Il y avait 10 participants par condition.

1.3 Procédure

Après avoir signé une feuille de consentement, le fonctionnement des trois tâches du MATB était expliqué aux participants. Suite à un entraînement de 10 minutes, les participants étaient informés de possibles défaillances du système automatisé avant d'effectuer le test de 90 minutes. Le test était composé de neuf sessions de 10 minutes. Dans la condition de système automatisé adaptatif, la quatrième et la septième sessions étaient manuelles. Les participants étaient informés de ces sessions manuelles par l'indication « *manual on* » s'affichant en bas de la tâche automatisée.

1.4 Mesures

La performance des participants a été mesurée dans chaque tâche : dans la tâche de surveillance, par le nombre de détections des participants lorsque le système ne fonctionnait pas ; dans la tâche de cible, par la proportion de temps du curseur dans la zone cible ; dans la tâche de gestion des ressources, par la moyenne des écarts absolues aux niveaux cibles (2500L), pour les deux réservoirs principaux (A et B).

Les données oculaires (temps cumulé de fixations) ont été mesurées par un *Eye-tracker* SMI (250 hz) en définissant des aires d'intérêts par rapport aux différentes tâches du MATB.

Nous avons également utilisé des questionnaires sur la confiance que les participants avaient envers eux-mêmes ou envers le système automatisé pour effectuer la tâche de surveillance.

2 RESULTATS

2.1 Données oculaires

L'analyse des données oculaires par le biais de comparaisons test-t entre la session 1 et la session 9 sur le temps total de fixation pour chaque niveau de fiabilité a permis d'étudier l'évolution du temps de regard sur la tâche automatisée.

1.1.1 Automatisation statique

Pour les niveaux 0%, les participants ont regardé la tâche automatisée significativement moins pendant la session 1 ($M=60.3$ s, $SE=10.6$ pour S1, $M=101.2$ s, $SE=10.0$ pour S9) que pendant la session 9 ($t(9) = -2,669$, $p<0.05$, $r = 0.44$). Pour les niveaux de fiabilité de 56,25% et 87,5%, aucune différence significative entre la session 1 et la session 9 n'a été enregistrée ($t(9) = 0,778$, $p = 0.45$ et $t(9)=1,210$, $p = 0.25$). Avec le niveau de fiabilité de 100%, les participants ont passé significativement plus de temps à regarder la tâche automatisée pendant la session 1 ($M=77.01$, $SE=16.7$ pour S1, $M=22.4$, $SE=5.4$ pour S9) que pendant la session 9 ($t(9) = 3,667$, $p<.01$, $r = .60$).

1.1.2 Automatisation adaptative

Le temps passé sur la tâche automatisée ne diffère pas entre la session 1 et 9 pour les niveaux 0%, 56.25% et 87.5 %. Cependant, une différence est observée pour le niveau 100%.

Table 1. Données descriptives du temps total de fixation et comparaisons t-test entre la session 1 et 9 pour chaque niveau de fiabilité.

Niveau de fiabilité	Session 1	Session 9	T-test
0 %	65.80s (10.5)	79.50s (12.3)	$t(9) = -1,433$, $p = 0.186$
56.25 %	88.7s (11.3)	108.2 (11.1)	$t(9) = -1,706$, $p = 0.122$
87.5 %	77.4s (10.8)	72.7s (15.5)	$t(9)=0,365$, $p = 0.725$
100%	76.1s (15.8)	36.3s (10.4)	$t(9) = 4,063$, $p<0.01$, $r = 0,64$

Note. Moyenne (erreur standard).

2.2 Performance

2.2.1. Automatisation statique

En ce qui concerne le niveau 0%, le taux de détection était plus faible dans la session 1 ($M=0.55$, $SE=0.13$) par rapport à la session 9 ($M=0.94$, $SE=0.02$), ($t(9) = -3,124$, $p<0,05$, $r = 0.52$). Pour

les niveaux de fiabilité de 56,25 % et 87,5 %, aucune différence significative entre la session 1 ($M=0.69$, $SE=0.05$ pour 56.25% et $M=0.65$, $SE=0.13$ pour le niveau 87.5) et la session 9 ($M=0.73$, $SE=0.11$ pour le niveau 56.25 et $M=0.75$, $SE=0.13$ pour le niveau 87.5) n'a été enregistrée ($t(9) = -0,439$, $p = 0.67$ et $t(9)=-1,0$, $p = 0.34$ respectivement).

2.2.2. Automatisation adaptative

Nous avons effectué des comparaisons de tests t entre chaque session manuelle et la session précédente (session 3/session 4 et session 6/session 7) ; chaque session manuelle et la session suivante (session 4/ session 5 et session 7/ session 8).

Avec un niveau de 0% du niveau de fiabilité, les participants ont eu un taux de détection plus faible dans la session 7 manuelle ($M=0.94$, $SE=0.02$) que dans la session 6 ($M=0.97$, $SE=0.01$), $t(9)=2,983$, $p<.05$. En ce qui concerne les niveaux 56,25% et 87,5%, les participants ont eu une meilleure performance dans la session 4 manuelle ($M=0.94$, $SE=0.03$ pour 56.25 et $M=0.79$, $SE=0.07$ pour le niveau 87.5) que dans la session 3 ($M=0.83$, $SE=0.06$ pour le niveau 56.25 et $M=0.50$, $SE=0.15$ pour le niveau 87.5), ($t(9)=-2,634$, $p<.05$ et $t(9)=-2,769$, $p<.05$ respectivement). Toutes les autres comparaisons n'étaient pas significatives.

2.3 Mesures subjectives

Une corrélation négative significative entre la confiance dans l'automatisation et le temps passé à regarder la tâche de surveillance a été constatée ($r(78)=-0,253$, $p<.05$). Plus la confiance dans l'automatisation diminue, plus le temps total de fixation sur la tâche de surveillance augmente.

3 DISCUSSION

Cette étude a permis de mettre en évidence qu'un niveau de fiabilité totalement parfait ou totalement imparfait n'engendrait pas les mêmes comportements visuels par les participants.

Plus précisément, nos résultats ont montré qu'une fiabilité élevée engendrait une baisse de l'attention visuelle accordée à la tâche automatisée. Bagheri et Jamieson (2004) avait trouvé les mêmes résultats. Avril et al. (2021) ont montré qu'en plus de la fiabilité, le type d'automatisation influençait également le comportement visuel des opérateurs. Notre étude montre également cette baisse d'attention visuelle déléguée à la tâche automatisée avec des niveaux élevés de fiabilité. Une telle baisse d'attention peut être vue comme une réponse naturelle et adéquate de l'opérateur humain qui fait de plus en plus confiance en un système automatisé lorsque celui-ci ne fait pas d'erreurs. Cependant, cette baisse d'attention peut mener à des accidents dans des situations où une baisse de la surveillance peut avoir des impacts irréversibles (nucléaire, aéronautique..). Cependant, Moray et Inagaki (2000) ont mis en garde contre une interprétation du comportement complaisant ou sceptique à l'égard de l'automatisation simplement basée sur un changement d'échantillonnage et une performance. Ils ont souligné l'importance de déterminer un comportement d'échantillonnage optimal et ont attribué la complaisance au sous-échantillonnage.

En ce qui concerne l'impact du système automatisé adaptatif, les sessions automatisées suivant la session manuelle de 10 minutes n'a pas été forcément suivi d'une différence significative du temps passé à examiner la tâche de surveillance lors des sessions automatisées suivantes (sauf pour le niveau de fiabilité de 100%). Cela montre que les sessions manuelles ont eu un impact sur le comportement visuel lors de la session automatisée suivante (Parasuraman et al., 1996).

Pour les sessions manuelles, le temps de fixation était le même pour les niveaux 56.25, 87.5 et 100%. Ce temps semble correspondre au temps nécessaire pour effectuer manuellement la tâche de

surveillance qui était auparavant automatisée, quel que soit le niveau de fiabilité de l'automatisation. En outre, ce temps est beaucoup plus long (26s) que le temps passé à examiner la tâche de surveillance dans le niveau de fiabilité de 0%. Cette différence correspond au coût visuel du retour au contrôle manuel pour les niveaux de fiabilité 87.5 et 100%. Cependant, si on s'intéresse au niveau de 56.25%, le temps passé à examiner la tâche automatisée lors de la première session manuelle est resté élevé lors des sessions automatisées suivantes. Cela indique qu'avec 56.25 % de fiabilité, le contrôle de la tâche automatisée a nécessité la même implication visuelle que qu'effectuer la tâche manuellement (Oakley et al., 2003 ; Onnasch et al., 2014 ; Wickens & Dixon, 2007)

4 BIBLIOGRAPHIE

- Avril, E., Valéry, B., Navarro, J., Wioland, L., & Cegarra, J. (2021). Effects of imperfect information and action automation on attentional allocation. *International journal of human computer interaction*. doi :10.1080/10447318.2020.1870817
- Bagheri, N., & Jamieson, G. A. (2004). Considering subjective trust and monitoring behavior in assessing automation-induced "complacency." *Human performance, situation awareness, and automation: Current research and trends*, 54–59.
- Cegarra, J., Valéry, B., Avril, E., Calmettes, C., & Navarro, J. (2020). OpenMATB: A Multi-Attribute Task Battery promoting task customization, software extensibility and experiment replicability. *Behavior Research Methods*, 1-11.
- Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human factors*, 37(2), 381-394.
- Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, 5(2), 113-153.
- Moray, N., & Inagaki, T. (2000). Attention and complacency. *Theoretical Issues in Ergonomics Science*, 1(4), 354-365. <https://doi.org/10.1080/14639220052399159>
- Navarro, J., & Osiurak, F. (2015). When do we use automatic tools rather than doing a task manually? Influence of automatic tool speed. *The American journal of psychology*, 128(1), 77-88.
- Oakley, B., Mouloua, M., & Hancock, P. (2003). Effects of automation reliability on human monitoring performance. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 47, p. 188–190). SAGE Publications Sage CA: Los Angeles, CA.
- Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human performance consequences of stages and levels of automation: An integrated meta-analysis. *Human factors*, 56(3), 476–488.
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human factors*, 52(3), 381–410.
- Parasuraman, R., Mouloua, M., & Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human factors*, 38(4), 665–679.
- Santiago-Espada, Y., Myer, R. R., Latorella, K. A., & Comstock Jr, J. R. (2011). The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide.
- Sheridan, T. B. (1970). On how often the supervisor should sample. *IEEE Transactions on Systems Science and Cybernetics*, 6(2), 140-145.
- Sheridan, T. B. (1987). Teleoperation, telepresence, and telerobotics: Research needs for space. *Human Factors in Automated and Robotic Space Systems. Proceedings, National Research Council, Washington, DC*, 279-291.
- Wiener, E. L. (1981, March). Complacency: Is the term useful for air safety. In *Proceedings of the 26th corporate aviation safety seminar* (Vol. 117, pp. 116-125).
- Wickens, C. D., & Dixon, S. R. (2007). The benefits of imperfect diagnostic automation: A synthesis of the literature. *Theoretical Issues in Ergonomics Science*, 8(3), 201–212.