[Page de garde]

HEC MONTRÉAL

L'impact des contre-mesures cognitives sur l'engagement et la résilience des opérateurs manufacturiers dans un contexte de travail assisté par l'IA

par
Loïc Couture

Pierre-Majorique Léger
HEC Montréal
Codirecteur de recherche

Robert Pellerin
Polytechnique Montréal
Codirecteur de recherche

Sciences de la gestion
(Spécialisation Expérience Utilisateur)

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences en gestion
(M. Sc.)

Août 2024

Résumé

Dans les usines intelligentes il est crucial de garder les opérateurs engagés et attentifs, que ce soit pour des raisons de sécurité ou pour garantir une gestion efficace des exceptions. De récentes études semblent indiquer que l'intégration d'intelligence artificielle (IA) dans les usines intelligentes peut réduire l'engagement des opérateurs, affectant ainsi leur capacité à réagir efficacement en cas d'exception. Dans ce mémoire, nous explorons l'utilisation de contre-mesures cognitives pour rehausser l'engagement et la performance des opérateurs dans des tâches manufacturières assistées par l'IA. Nous posons l'hypothèse que si les opérateurs sont plus engagés dans leur travail grâce aux contre-mesures cognitives, alors ils seront plus performants, plus motivés et réagiront mieux lors de la gestion d'exceptions. Deux contre-mesures cognitives ont été étudiées: la réalité augmentée (RA) et les systèmes de rétroaction du niveau d'engagement (SRNE). Dans ce mémoire, nous proposons d'abord un nouveau SRNE adapté au milieu manufacturier qui utilise des mesures physiologiques faciles à collecter en mouvement telles que la respiration et l'accélération. En second lieu, nous évaluons l'effet des deux contre-mesures cognitives (la RA et le SRNE développé) sur l'engagement, la motivation et la performance des opérateurs en contexte de travail assisté par l'IA, ainsi que sur leur résilience en cas d'échec d'automatisation. Les résultats suggèrent que le SRNE conçu était capable de prédire les niveaux d'engagement de notre banque de données d'entraînement avec une précision de 80.95%. Les résultats indiquent également que les contre-mesures aident les opérateurs à développer des compétences plus résilientes. Les travailleurs utilisant les contre-mesures ont montré une plus faible réduction de précision lorsque l'automatisation a été retirée comparativement au groupe contrôle. De plus, les résultats indiquent que le SRNE peut améliorer l'engagement comportemental des opérateurs lors du travail assisté par l'IA. En effet, le SRNE a mené à une plus grande accélération dans la tâche assistée par l'IA. Toutefois, contrairement à nos attentes, les contre-mesures n'ont pas eu d'impacts sur la motivation, l'engagement cognitif et l'engagement émotionnel. Nos résultats montrent qu'il est possible d'utiliser les contre-mesures cognitives pour atténuer certains problèmes de performance liés à l'automatisation.

Mots clés : Contre-mesure cognitive, Réalité augmentée, Rétroaction du niveau d'engagement, Résilience, Engagement, Motivation, Fabrication, Intelligence artificielle

Méthodes de recherche : Expérimentation, Science de design

Abstract

In smart factories, it is crucial to keep operators engaged and attentive, whether for safety reasons or to ensure effective management of exceptions by the operators. Recent studies suggest that the integration of artificial intelligence (AI) in smart factories may reduce operator engagement, thereby affecting their ability to respond effectively in case of an exception. In this thesis, we explore the use of cognitive countermeasures to enhance operator engagement and performance in AI-assisted manufacturing tasks. We hypothesize that if operators are more engaged in their work due to cognitive countermeasures, they will be more performant, more motivated, and better equipped to handle exceptions. Two cognitive countermeasures were studied: augmented reality (AR) and a real-time engagement level feedback systems (RTELFS). In this thesis, we first propose a new RTELFS adapted to the manufacturing environment, which uses physiological measures that are easy to collect while in motion, such as respiration and acceleration. Secondly, we evaluate the effect of the two cognitive countermeasures (AR and the developed RTELFS) on operator engagement, motivation, and performance in the context of AI-assisted work, as well as their resilience in the event of automation failure. The results suggest that the designed RTELFS was capable of predicting engagement levels from our training data with an accuracy of 80.95%. The results also indicate that the countermeasures help operators develop more resilient skills. Workers using the countermeasures showed a smaller reduction in accuracy when automation was removed compared to the control group. Additionally, the results indicate that the RTELFS can improve behavioral engagement during AI-assisted work. Indeed, the RTELFS led to a greater acceleration in the AI-assisted task. However, contrary to our expectations, the countermeasures did not impact motivation, cognitive engagement, or emotional engagement. Our findings demonstrate that it is possible to use cognitive countermeasures to mitigate certain performance issues related to automation.

Keywords: Cognitive Countermeasure, Augmented Reality, Engagement Feedback, Resilience, Engagement, Motivation, Manufacturing, Artificial Intelligence

Research methods: Experimentation, Design Science

Table des matières

Liste des tableaux

Liste de figures

Chapitre 2

Chapitre 3

Liste des abréviations

(FR)

EEG : Électroencéphalographie

IA : Intelligence artificielle

RA : Réalité augmentée

SRNE : Système de rétroaction du niveau d'engagement

(EN)

AI: Artificial intelligence

AR: Augmented Reality

EEG: Electroencephalography

fNIRS: Functionnal near infrared spectroscopy

RTELFS: Real-time engagement-level feedback system

UWES: Utrecht work engagement scale

TS-EI: Task-specific engagement index

HRV: Heart rate variability

HF-HRV: High frequency band power of heart rate variability

LF-HRV: Low-frequency band power of heart rate variability

RMSSD: Root mean square of successive RR interval differences

ANS: Autonomous nervous system

LC-NE: locus-ceruleus norepinephrine system

RR: Respiration rate

EDA: Electrodermal activity

LF/HF: Ratio of low frequency to high frequency power of HRV

Avant-propos

Ce travail a été effectué dans le cadre d'une maîtrise en expérience utilisateur à HEC Montréal. Le sujet de mémoire a été enregistré et approuvé par l'administration du programme et une autorisation a été reçue pour rédiger ce mémoire par article. L'inclusion des deux articles dans ce mémoire a été approuvée par tous les co-auteurs. Le premier article (chapitre 2) a été publié dans le journal *Sensors and Transducers* en mai 2024. L'article publié est une version étendue d'un article de conférence qui avait été rédigé en préparation pour la conférence *ARCI2024 (Automation, Robotics and Communication for Industry 4.0/5.0)* qui a eu lieu du 7 au 9 février 2024. L'article de conférence présenté à cette conférence se trouve en annexe A. Le second article (chapitre 3), a été rédigé pour soumission au journal *International Journal of Production Research (IJPR).* Au moment de la remise de ce mémoire, l'article n'a pas été soumis au journal.

La recherche présentée dans ce mémoire a été approuvée par le comité d'éthique de la recherche (CER) de HEC Montréal en mai 2023 sous le numéro de projet 2023-5427. Conformément aux exigences, le projet a été complété de manière éthique.

Remerciements

Je tiens d'abord à remercier mes deux co-directeurs de recherche, Pierre-Majorique Léger et Robert Pellerin, pour leur aide et leur mentorat tout au long de ce projet. Votre soutien a été pour moi une véritable source d'inspiration. Pierre-Majorique, je tiens spécialement à te remercier pour la confiance que tu as placée en moi durant les dernières années et pour les opportunités que j'ai pu vivre en travaillant sous ta direction. Ces expériences m'ont permis de me développer tant sur le plan professionnel que personnel et je t'en serai toujours reconnaissant. Robert, je te remercie pour les belles expériences et surtout pour les amitiés que j'ai pu développer à travers nos collaborations.

Ensuite, je tiens à remercier Mario Passalacqua pour son mentorat et son soutien tout au long de ce travail. Merci d'avoir pris le temps de m'aider dans la collecte de données, la rédaction de mes articles, ainsi que dans mes réflexions. Tu as réellement été un mentor exceptionnel.

Je souhaite également remercier mes collaborateurs Laurent Joblot et Florian Magnani pour leur accueil et leur aide lors de la collecte de données. J'espère avoir la chance de poursuivre nos collaborations dans le futur.

Je tiens également à remercier ma copine Frédérique pour son soutien et ses encouragements tout au long de la réalisation de ce travail, ainsi que mes parents, qui ont toujours été présents pour m'aider à pousser mes réflexions.

Enfin, je tiens à remercier tout le personnel du Tech3Lab pour leur soutien technique et pour avoir créé une communauté scientifique qui a enrichi mon expérience académique et qui a su me guider dans toutes les étapes de la réalisation de ce projet.

Je tiens également à souligner la contribution de nos partenaires CRSNG, Prompt et IVADO, qui ont permis de financer ce projet de recherche.

Chapitre 1
Introduction

1.1 Mise en contexte

Au cours des dernières années, nous avons observé une croissance du déploiement des systèmes d'intelligence artificielle (IA) dans les usines intelligentes. Les capacités de traitement de l'information et de projection des systèmes d'IA permettent aux systèmes de devenir de plus en plus autonomes, leur permettant ainsi de fonctionner au-delà des heures de travail et entraînant des augmentations significatives de la productivité des systèmes (Yang et al., 2021). De plus, l'intégration d'IA a aussi démontré des effets positifs sur la productivité des travailleurs (Plathottam et al., 2023; Raj and Seamans, 2018). En effet, ces systèmes peuvent être utilisés pour réduire la charge de travail des opérateurs en prenant en charge certaines tâches répétitives ou à faible valeur ajoutée (Tortorella et al., 2024), ou encore pour augmenter les capacités des travailleurs, par exemple en leur fournissant des instructions en temps réel (Sahu, Young and Rai, 2021).

Actuellement, les principales utilisations de l'IA dans les usines intelligentes incluent la détection de défauts, le suivi de la maintenance prédictive, la gestion des coûts et de l'énergie, ainsi que le développement de robots et de systèmes de conduite autonome (Nti et al., 2022), mais de nouvelles applications émergent progressivement. Par exemple, Mypati et al. (2023) ont récemment proposé plusieurs applications d'IA dans les domaines du moulage, du formage et de la finition dans les fonderies. Manikandan et al. (2023) ont souligné le besoin de développer des outils d'IA dans les procédés avancés d'usinage ainsi que dans les techniques de soudage des métaux. De plus, He et al. (2023) ainsi que Mattera, Nele et Paolella (2024) ont récemment proposé de nouvelles façons d'utiliser l'IA dans les méthodes de fabrication additives. Avec des applications de plus en plus diversifiées, on s'attend à ce que l'IA devienne de plus en plus omniprésente dans les usines intelligentes du futur.

Lorsque les opérateurs manufacturiers travaillent avec des systèmes d'IA, il devient crucial pour eux de maintenir des niveaux élevés d'engagement et d'attention dans leur travail. En effet, Mangler et al. (2021) expliquent que les opérateurs dans les usines intelligentes gèrent souvent une variété de sous-systèmes et interagissent avec plusieurs niveaux d'automatisation, ce qui peut augmenter la charge cognitive requise pour réaliser leur travail (Yamamoto, 2019). Cette demande mentale accrue peut conduire à des erreurs, en particulier lorsque les opérateurs sont inattentifs (Mangler et al., 2021), ce qui peut potentiellement compromettre la qualité de la production (Yung et al., 2020) ou, dans les cas les plus graves, entraîner des accidents (Naderpour, Nazir et Lu, 2015). De plus, lors de l'intégration de nouvelles technologies manufacturières comme l'IA, les entreprises s'attendent à ce que les opérateurs soient capables de rapidement identifier les erreurs des systèmes et puissent intervenir efficacement en cas de problèmes d'automatisation (Endsley and Kiris, 1995). En cas de défaillance de l'automatisation (aussi appelé exception) il devient particulièrement important que les opérateurs humains puissent reprendre le processus de manière efficace tout en maintenant un état calme afin de limiter les pertes de productivité possibles (Romero et Stahre, 2021).

Il y a cependant une inquiétude croissante que l'intégration d'intelligence artificielle en milieu manufacturier puisse avoir des effets négatifs sur l'engagement et la motivation des opérateurs, affectant ainsi leur capacité à superviser adéquatement des systèmes automatisés ainsi qu'à répondre efficacement aux problèmes d'automatisation (Endsley, 2023). Cette inquiétude est exacerbée par les résultats d'une récente étude par Passalacqua et al. (2024) dans laquelle les chercheurs démontrent que l'utilisation d'une IA très performante lors de la formation d'opérateurs manufacturier avait réduit leur engagement, leur motivation ainsi que leur performance lorsque l'IA a été retirée, comparativement aux opérateurs qui s'étaient formés avec un système moins performant, nécessitant des interventions humaines. Ce résultat est plutôt alarmant puisque les applications de l'IA sont de plus en plus performantes, variées et de plus en plus intégrées dans les processus manufacturiers des entreprises (Li et al., 2017). Il semble donc impératif de trouver des solutions pour s'assurer que l'intégration d'IA puisse promouvoir la performance et le bien-être des opérateurs humains au lieu de les dégrader.

Les contre-mesures cognitives représentent des solutions très prometteuses pour rehausser l'engagement et la performance des opérateurs dans des contextes hautement automatisés. Les contre-mesures cognitives sont définies de manière globale comme étant des stratégies, des techniques ou des outils qui permettent d'améliorer ou de maintenir des performances cognitives durant une tâche (Dehais et al., 2010). Elles sont souvent utilisées pour limiter certains biais cognitifs comme la baisse d’engagement (Karran et al., 2019; Demazure et al., 2021) ou le « tunnel cognitif », un état cognitif où les opérateurs adoptent un focus intense sur certaines informations, négligeant les informations externes (Dehais et al., 2010). Bien que la littérature scientifique bénéficierait d’une définition plus spécifique des contre-mesures cognitives, dans ce travail nous faisons référence à ce concept pour désigner des outils et des méthodes pour assurer que les opérateurs puissent maintenir des niveaux optimaux d’engagement dans leur travail et éviter les distractions.

Un bon exemple de contre-mesure cognitive a été développé par Dehais, Causse et Tremblay (2011) pour réduire les baisses de performances dues au « tunnel cognitif ». La contre-mesure développée par ces chercheurs consistait à retirer momentanément les informations principales du tableau de bord des opérateurs afin d’afficher les signaux externes. Ceci forçait les opérateurs à prendre connaissance des signaux externes même en état de « tunnel cognitif » ce qui a mené à une meilleure prise de décision globale. Ce système représente un bon exemple de contre-mesure cognitive, mais n’est pas applicable pour rehausser l’engagement des opérateurs. Pour rehausser l’engagement, deux contre-mesures cognitives semblent particulièrement prometteuses : les systèmes de rétroaction du niveau d’engagement (SRNE) et la réalité augmentée (RA).

Les SRNE sont des systèmes qui utilisent des réponses physiologiques (i.e., électroencéphalographie, variabilité de la fréquence cardiaque) et des métriques de performance pour mesurer les niveaux d'engagement cognitif des opérateurs en temps réel. Ces systèmes affichent ensuite le niveau d’engagement mesuré en temps réel, permettant aux opérateurs de rester conscients de leur niveau d’engagement dans la tâche. Ces systèmes agissent comme des contre-mesures cognitives en permettant aux opérateurs de s’ajuster en temps réel face à des baisses d’engagement afin de maintenir des niveaux d’engagement soutenus dans leur travail. Un exemple notable de SRNE a été développé par Demazure et al. (2021) et consistait à changer la couleur de fond d’un logiciel de gestion de ressources (ERP) en fonction du niveau d’attention soutenue des opérateurs mesuré à partir d’un index d’électroencéphalographie (EEG). Ce système a démontré un grand potentiel pour aider les opérateurs à maintenir des niveaux d’attention soutenue optimaux dans une longue tâche de suivi passif de processus d’entreprise (Karran et al., 2019).

La réalité augmentée (RA) démontre également un fort potentiel pour renforcer l'engagement des opérateurs en milieu manufacturier. En permettant de manipuler les informations accessibles aux opérateurs, la RA offre une capacité significative à orienter leur attention vers les données les plus pertinentes, en faisant ainsi un outil puissant à utiliser comme contre-mesure cognitive. En milieu manufacturier, la RA a été beaucoup utilisée pour afficher des instructions d'assemblage aux opérateurs en temps réel et directement sur leur environnement de travail (Wang et al., 2022; Werrlich et al., 2017). Dans ce contexte, la RA est principalement utilisée pour améliorer l'accessibilité des instructions d'assemblage, ce qui réduit la charge cognitive liée à la recherche d'informations. Elle permet également de limiter les pertes de productivité et les distractions potentielles associées au passage de la tâche principale à la consultation d'un manuel d'instructions ou d'un tableau de bord. En ce qui concerne l'effet de la RA sur l'engagement des opérateurs, Nguyen et Meixner (2019) ont démontré que l'utilisation de ludification (*gamification*) avec la RA dans des tâches d'assemblage avait globalement rehaussé l'engagement des opérateurs manufacturiers. De plus, Lam et al. (2021) ont démontré que l'entraînement d'opérateurs manufacturiers à l'aide de RA permettait d'améliorer la rétention d'information comparativement à l'entraînement avec papier, ce qui semble suggérer un plus grand engagement lors de l'entraînement.

Malgré le fort potentiel de ces deux contre-mesures, actuellement, elles n'ont jamais été étudiées en laboratoire comme moyen pour mitiger les impacts négatifs de l'IA sur l'engagement et la performance des opérateurs manufacturiers. De plus, les SRNE présentés dans la littérature utilisent principalement l'EEG pour mesurer l'engagement des opérateurs, une méthode peu adaptée au contexte manufacturier où les opérateurs sont souvent en mouvement.

1.2 Objectifs et questions de recherche

Ainsi, l'objectif de ce travail est double. Premièrement, nous souhaitons développer un système de rétroaction du niveau d'engagement qui serait adapté au contexte manufacturier. Ensuite, nous souhaitons évaluer l'effet des deux contre-mesures cognitives (c'est-à-dire de la RA et le système de rétroaction du niveau d'engagement) sur l'engagement, la motivation, la performance et la résilience des opérateurs manufacturiers en contexte de travail assisté par l'IA. Les questions de recherches qui ont guidé notre travail sont donc les suivantes :

1. Comment pouvons-nous concevoir un système de rétroaction du niveau d'engagement adapté au milieu manufacturier ?
2. Quel est l'effet des deux contre-mesures cognitives (i.e., la RA et le système de rétroaction du niveau d'engagement) sur l'engagement, la performance, la motivation et la résilience des opérateurs manufacturiers dans un environnement de travail assisté par l'IA?

Nous posons l'hypothèse que si les opérateurs sont plus engagés dans leur travail, alors ils seront plus motivés et performants lors du travail assisté, et développeront également des compétences plus résilientes leur permettant de mieux réagir lors de situations critiques, par exemple lorsque l'automatisation échoue.

1.3 Structure du mémoire

Ce mémoire est structuré en quatre chapitres. Le présent chapitre offre une mise en contexte générale et introduit les principaux objectifs de recherche qui ont guidé ce travail. Le chapitre 2 répond à la première question de recherche dans un article scientifique détaillant le processus de conception d'un système de rétroaction du niveau d'engagement adapté au contexte manufacturier. Dans le chapitre 3, nous adressons la deuxième question de recherche dans un article portant sur l'effet de la RA et du système de rétroaction du niveau d'engagement développé dans le chapitre 2 sur l'engagement, la motivation, la performance et la résilience des opérateurs manufacturiers dans un environnement de travail assisté par l'IA. Enfin, le chapitre 4 contient une conclusion qui résume les principaux résultats obtenus dans ce mémoire, les contributions majeures, ainsi que les perspectives pour de futurs projets de recherche.

La structure des articles scientifiques des chapitres 2 et 3 a été adaptée pour publication dans des journaux académiques, ce qui explique pourquoi ceux-ci sont rédigés en anglais. Une version de l'article présenté dans le chapitre 2 a été publiée en mai 2024 dans le journal *Sensors and Transducers* (Couture et al., 2024a). Additionnellement, une version plus courte de cet article a été présentée à la conférence *ARCI2024* en février 2024 (Couture et al., 2024b). Cet acte de conférence se trouve en annexe. L'article du chapitre 3 a été rédigé en vue d'une publication dans le journal *International Journal of Production Research (IJPR)*, mais n'a pas encore été soumis.

1.4 Contribution

Tableau 1.1 Tableau de contribution

Étapes	**Contribution**
Questions de recherche	Identifier les “gaps” dans la littérature et le problème de recherche **– 85%** - Identifier les écarts dans la littérature **- 100%** - Définir les questions de recherches **- 80%** - Définir les hypothèses **– 80%** - Définir les construits à tester **– 80%** - Les directeurs de recherche ont aidé à guider les questions de recherche et hypothèse
Revue de littérature	Rechercher et lire les articles pertinents pour ce projet de recherche **– 100%** - Méthodologie de recherche d’articles **– 100%** - Lecture des articles **– 100%**
Design expérimental	Application au CER **– 100%** *(Chapitre 2)* Protocole expérimental **– 100%** Setup expérimental **– 95%** Stimuli expérimentaux **– 75%** - Le système testé a été entièrement développé (codé) par l’étudiant **- 75%** - L’index d’engagement a été calculé par le statisticien du Tech3lab ce qui explique la contribution de 75% de l’étudiant. *(Chapitre 3)* Protocole expérimental **– 75%** - Un co-auteur ayant réalisé une expérimentation très similaire a beaucoup aidé l’étudiant à définir le protocole expérimental, ce qui explique la contribution de l’étudiant de 75% Setup expérimental **– 0%** - Cette expérience a été réalisée dans une usine-école en France. Ce sont des collaborateurs de cet établissement qui ont créé la simulation manufacturière. Stimuli expérimentaux **– 50%** - Système de rétroaction du niveau d’engagement – 75% - Le système de réalité augmentée a été réalisé par des collaborateurs – 0%
Recrutement	Recrutement **– 75%** - Configuration de l’outil de recrutement **– 100%** - Recrutement et compensation (chapitre 2) **– 100%** - Recrutement et compensation (chapitre 3) **– 50%** - Un collaborateur a fortement aidé dans le recrutement de participants pour l’expérimentation du chapitre 3, mais la compensation a entièrement et prise en charge par l’étudiant. Ceci explique la contribution de 50% de l’étudiant.
Collecte de données	Collecter les données et superviser les opérations **– 100%** - Collecte de données (chapitre 2) **– 100%** - Collecte de données (chapitre 3) **– 100%**

Analyse	Extraction et synchronisation des données **– 85%** - Extraction et synchronisation des données (chapitre 2) **– 100%** - Extraction et synchronisation des données (chapitre 3) **– 100%** - Calcul de moyennes à partir des données brutes (chapitre 2) **– 100%** - Calcul de moyennes à partir des données brutes du chapitre 3 a été réalisé par le statisticien du Tech3Lab, mais l'étudiant a fourni un guide pour organiser et comprendre les données **– 25%** Analyse statistique **– 90%** - Analyses statistiques (chapitre 2) **– 100%** - Analyses statistiques (chapitre 3) **– 80%** - Le statisticien du Tech3Lab a fourni un guide à l'étudiant sur les tests statistiques recommandés pour chaque variable. Ceux-ci ont été réalisés entièrement par l'étudiant avec le logiciel SAS Création de tableaux et graphiques **– 100%**
Rédaction	Rédaction **– 95%** - Les co-auteurs ont aidé à réviser et à fournir des commentaires pour les deux articles

Chapitre 2

Adaptive System to Enhance Operator Engagement During Smart Manufacturing Work*

Loïc Couture[1], Mario Passlacqua[2], Laurent Joblot[3], Florian Magnani[4], Robert Pellerin[2] and Pierre-Majorique Léger[1]

[1] HEC Montréal, Tech3Lab, 3000 Chemin de la Côte-Sainte-Catherine, H3T 2A7, Canada

[2] Polytechnique Montréal, 2500 Chemin de Polytechnique, H3T 1J4, Canada

[3] Arts et Métiers Institute of Technology, LISPEN / HESAM University / UBFC, France

[4] Université Aix-Marseille, CERGAM, Centrale Méditerranée, 38, rue Frédéric Joliot-Curie, 13013, France

Abstract: Sustaining optimal task engagement is becoming vital in smart factories, where manufacturing operators' roles are increasingly shifting from hands-on machinery tasks to supervising complex automated systems. However, because monitoring tasks are inherently less engaging than manual operation tasks, operators may have a growing difficulty in keeping the optimal levels of engagement required to detect system errors in highly automated environments. Addressing this issue, we created an adaptive task engagement feedback system designed to enhance manufacturing operators' engagement while working with highly automated systems. Utilizing real-time acceleration, heart rate, and respiration rate data, our system provides an intuitive visual representation of an operator's engagement level through a color gradient, ensuring operators can stay informed of their engagement levels in real-time and make prompt adjustments if required. This paper elaborates on the six-step process that guided the development of this adaptive feedback system. We developed a task engagement index by leveraging the physiological distinctions between more and less engaging manufacturing scenarios and using automation to induce lower engagement. This index demonstrates a prediction accuracy rate of 80.95% for engagement levels, as demonstrated by a logistic regression model employing leave-one-out cross-

* Une version de cet article a été publiée dans le journal *Sensors and Transducers* en mai 2024 (Couture et al., 2024a)

validation. The implications of deploying this adaptive system include enhanced operator engagement, higher productivity and improved safety measures.

Keywords: Engagement, Adaptive system, Manufacturing, Industry 5.0, Human-machine interaction, Design science.

2.1 Introduction

Recent advances in manufacturing technologies have significantly expanded the capabilities of automation, enabling even traditionally human-centric tasks to be automated. When automating such tasks, we frequently see the role of human operators transitioning from manual labor to supervisory roles, which can have negative implications for operators' engagement in their work (Passalacqua et al., 2024a). In the context of Industry 5.0, which emphasizes the importance of workers' interests and well-being, ensuring that operators remain engaged and stimulated in their roles becomes crucial (Lu et al., 2022; Goujon et al., 2024). This approach is not only fundamental to their personal development but is also imperative for enhancing their decision-making skills, especially in increasingly complex work environments (Goujon et al., 2024; Rosin et al., 2021; Rosin et al., 2022). By prioritizing the engagement and stimulation of operators, organizations can navigate the challenges of modern manufacturing landscapes more effectively, ensuring that technological advancements contribute positively to the work experience of human operators (Passalacqua et al., 2024b). Engagement varies in definition across the literature. It is often used either as “task engagement” or "operator engagement" to describe the effective allocation of attentional resources towards the task objectives (Pope et al., 1995; Matthews et al., 2002; Dehais et al., 2020). This definition focuses on the cognitive aspects of the worker experience. Alternatively, terms like “work engagement” and “employee engagement” are used to characterize a positive, fulfilling psychological state related to work (Mazzetti et al., 2021; Bakker and Demerouti, 2008; Hallberg and Schaufeli, 2006; Saks, 2006) that encompasses the cognitive, behavioral, and emotional aspects of the work experience. Given the prevalent focus on the cognitive dimension of engagement in existing research on human-machine interaction, we employ ‘task engagement’ to represent the cognitive aspect of the work experience. This decision acknowledges the broader spectrum of engagement but aligns our focus with the extensive body of research emphasizing cognitive engagement in the context of human-machine interaction.

To effectively oversee automated systems, an operator must remain alert to various signals, referred to as arousal, and must stay focused on the task at hand, known as task engagement (Dehais et al., 2020). There is a sweet spot of arousal and task engagement that operators need to sustain to ensure adequate and optimal monitoring (Dehais et al., 2020). If an operator is disengaged, they risk becoming distracted with mind-wandering (Cheyne et al., 2009), whereas being overly engaged can lead to tunnel vision, preventing the operator from staying alert to external signals (Pooladvand and Hasanzadeh, 2023). Similarly, if an operator has too high or too low arousal, it might affect their cognitive capabilities (Dehais et al., 2020). However, it can be challenging for operators to maintain an optimal level of engagement in monitoring tasks, mainly because monitoring tasks are generally less engaging than manual operation tasks (Passalacqua et al., 2024a). Consequently, an under-stimulated operator is much more likely to be distracted, which reduces their ability to detect system errors (Thomson et al., 2015; Smallwood and Schooler, 2006). This monitoring difficulty increases as the level of automation rises (Parasuraman, 2000). Therefore, in increasingly intelligent factories, there may be a growing difficulty in detecting errors in automated systems.

To limit these performance declines, one method is to ensure that operators can maintain optimal levels of task engagement during their work (Couture et al., 2024; Yurish, 2024, p.232). The work of Karran et al. (2019) is particularly promising in this regard. Their research demonstrated the potential of using real-time engagement level feedback to improve users' attentiveness during a passive monitoring task. In their paper, they used an adaptive system developed by Demazure et al. (2021) that continuously informed operators of their level of engagement in the task through a color gradient, using electroencephalography (EEG) measurements to infer engagement. While this solution has shown promising results, it faces significant limitations in a manufacturing context, primarily due to the high sensitivity of EEG to movement. Therefore, our study seeks to adapt this approach for manufacturing, aiming to develop a tool designed to help manufacturing operators maintain optimal engagement levels when working with highly automated systems. The primary aim of this adaptation is to leverage engagement metrics collectible from mobile operators. The research question that guided the system's design is: How can the engagement feedback system proposed by Demazure et al. (2021) be effectively adapted and implemented in a manufacturing setting to monitor and enhance the engagement of mobile manufacturing operators?

The structure of this paper is outlined as follows. Section 2 contains an overview of the current solutions to enhance operator engagement, why we hypothesize that adaptive feedback systems represent a good solution, and which methods are used to measure task engagement in the literature. Section 3 contains the research objectives that guided our design. In section 4, we detail the six-step process that led to developing this new innovative feedback system. The results we obtained during the design process are detailed in Section 5. In section 6, we discuss the results, and in section 7, we provide our concluding remarks along with limitations of the system and insights into future developments.

2.2 Background

2.2.1 Solutions to Enhance Task Engagement During Monitoring Tasks

Solutions to keep operators cognitively engaged during monitoring tasks can be categorized into multi-tasking and adaptive interfaces. Multi-tasking involves engaging the operator with non-task-related tasks when they experience disengagement. Naujoks et al. (2018) showed that engaging in secondary tasks reduced the reaction time of drivers when they needed to regain control, and Atchley et al. (2011) showed that talking while driving could improve driving performance. However, one limitation of multi-tasking is that it requires the operator to divert some of their attention to a secondary task, which diminishes the total level of engagement they can apply to the primary task (Argyle et al., 2021). For this reason, adaptive systems appear to be a better alternative for keeping operators engaged when monitoring systems. Adaptive systems infer the cognitive state of human operators and use this information to adapt in real time. According to Hinss et al. (2022), there are two main types of adaptive systems: adaptive automation and adaptive interfaces. Adaptive automation allows for the dynamic adjustment of task allocation based on the cognitive state of operators (Scerbo, 2007; Feigh et al., 2012). The purpose of these systems is to reduce the level of automation of automated systems when decreases in engagement are detected. This decrease in automation necessitates that the operator takes on more stimulating tasks, thereby potentially restoring their engagement to a level considered adequate for environments characterized by higher levels of automation.

The second type of adaptive system consists of adaptive interfaces. Feigh et al. (2012) developed a taxonomy for adaptive interfaces, proposing four modalities of adaptation, including task allocation, which refers to adaptive automation. The three other modalities are the following. When operators are in a state of cognitive disengagement adaptive interfaces can adjust task prioritization, for example, by requesting the operator to perform tasks that are either more stimulating or require less engagement. They can also adapt the interaction between the operator and the system, for instance, by changing the layout of components or the mode of interaction (e.g., from haptic to vocal). Lastly, the content of the interface can be adapted, for example, by increasing the amount of information displayed when the operator is engaged and reducing it when they are less engaged. **Figure 2.1** summarizes the different solutions to keep operators engaged based on the works of Feigh et al. (2012) and Hinss et al. (2022).

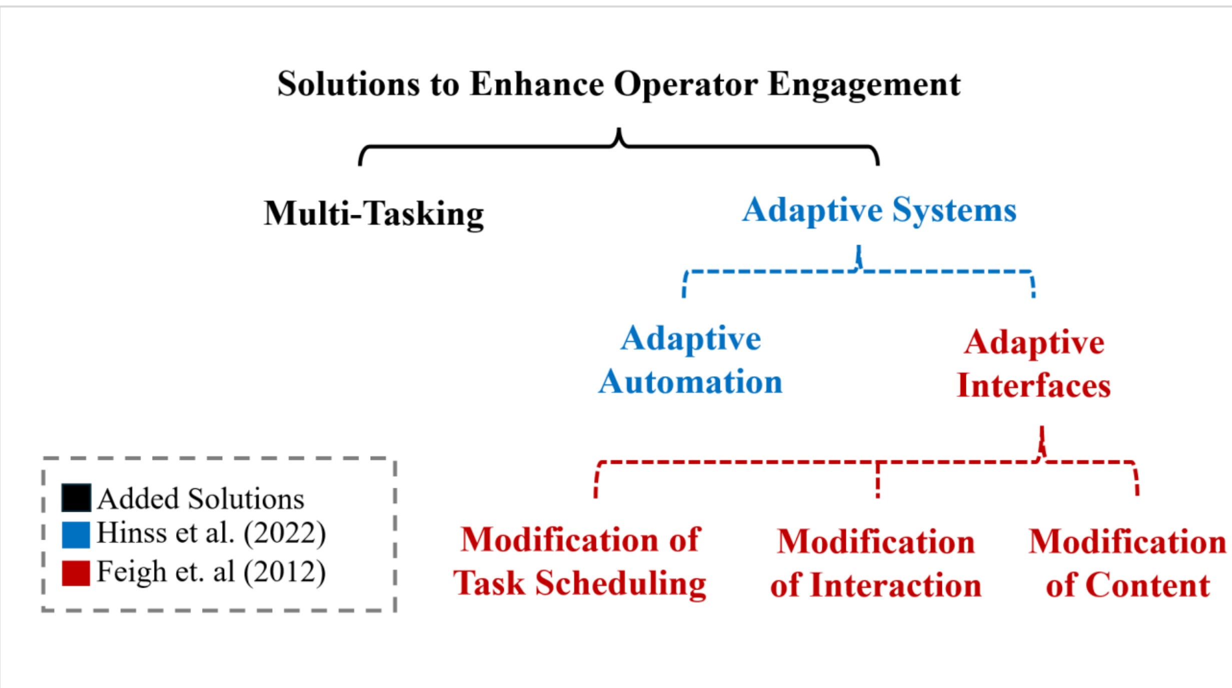

Figure 2.1 Categorization of current solutions to enhance operator engagement during surveillance work

Despite the introduction of adaptive automation and adaptive interfaces years ago, very few adaptive systems with generalizable applications have been developed (Bernabei and Costantino, 2024). The reasons for this include the absence of comprehensive multimodal models to infer operators' cognitive states (Bernabei and Costantino, 2024), the dependence of adaptive systems applications on the specific working environment in which they are developed, and constraints regarding the physiological data collection across different work environments. However, feedback systems and alarm systems distinguish themselves from other adaptive systems solutions by providing a passive solution that does not need to interface with various systems and introduces minimal distraction, making it relatively easy to apply across different contexts. These systems would fall into the adaptive interfaces category by modifying the content of the interface (e.g., the presence or absence of visual or auditory cues). The key distinction between feedback and alert systems lies in the way they present countermeasures. Adaptive feedback systems usually give continuous feedback to operators on their cognitive and emotional states, whereas alert systems typically wait for specific thresholds to be reached before notifying the operator. Karran et al. (2019) compared these two approaches, revealing that the continuous display of mental state had a greater impact on operator engagement compared to displaying the mental state after a disengagement threshold was reached. Therefore, we opted for the development of an adaptive feedback system.

The feedback system developed by Demazure et al. (2021) offers a promising approach to enhancing task engagement during monitoring tasks by providing operators with real-time feedback on their level of engagement. This innovation keeps operators continuously aware of their engagement, promoting immediate adjustments as needed. Currently, the manufacturing sector lacks such adaptive feedback systems specifically designed for engagement levels. Despite this, the broad applicability and proven effectiveness of engagement-level feedback systems underscore their potential value in maintaining operator engagement, particularly in environments requiring high task engagement.

2.2.2 Measuring Task Engagement

Task engagement, or the cognitive aspect of work engagement, is most commonly measured using questionnaires or observational metrics. The most commonly used questionnaire to measure engagement is the Utrecht Work Engagement Scale (UWES) (Schaufeli et al., 2003). The UWES questionnaire facilitates the measurement of the multi-dimensional concept of engagement, defined earlier. It encompasses three dimensions: cognitive engagement in work (related to the absorption dimension of the questionnaire), behavioral engagement at work (related to the vigor dimension of the questionnaire), and emotional engagement (related to the dedication dimension of the questionnaire). Given that the absorption dimension encapsulates the cognitive aspect of engagement, it can be used as a measure of task engagement. Regarding observational metrics, task performance metrics are the most commonly utilized measures. Performance-based measures of task engagement include, for example, error detection performance (Passalacqua et al., 2024a; Parasuraman et al., 1993), sampling time (Moray and Inagaki, 20000), and reaction time (Körber et al., 2015). While performance-based and subjective metrics effectively identify instances of lower operator engagement when monitoring automated systems, both these measures have their limitations. Questionnaires, which depend on post-task subjective assessments, are prone to biases such as recall bias (de Guinea et al., 2014). Performance metrics, while serving as useful engagement proxies, do not directly measure engagement and can be influenced by various extraneous factors. A solution to complement the limitations of questionnaires and performance metrics is the use of physiological measures, which allow for the continuous measurement of the participant's state throughout the task, without interference, thus limiting biases (de Guinea et al., 2013). Consequently, recent research has increasingly focused on physiological metrics to understand the impact of operators' mental states on performance, providing deeper insights into engagement dynamics (de Guinea et al., 2014; Riedl et al., 2020; Passalacqua et al., 2020). The physiological measures used to infer task engagement include eye-tracking, electroencephalography (EEG), heart rate variability (HRV), respiration rate (RR), electrodermal activity (EDA), and functional near-infrared spectroscopy (fNIRS). These various modalities and their application in measuring task engagement are presented in this section.

Eye-tracking

Eye-tracking tools detect where an operator's gaze lands and measure pupil diameter, both indicators of task engagement (Vasseur et al., 2023). An operator is considered engaged when their gaze is on the main points of interest of a task and disengaged when their gaze deviates from them. When analyzing the gaze of operators, we typically distinguish between fixations and saccades. Fixations refer to the moments when the eyes are relatively stationary and are focused on a specific point for a period of time, generally lasting between 180 and 330 milliseconds (Carter and Luke, 2020), while saccades are rapid eye movements between fixations. One way to interpret task engagement using gaze data is to use the position, frequency, and duration of fixations (Gouraud et al., 2018). Pupil diameter is used as an indicator of task engagement and cognitive fatigue (Hopstaken et al., 2015) because this physiological mechanism is controlled by the locus ceruleus norepinephrine system (LC-NE) region of the central nervous system, which is also responsible for regulating attention (Benarroch, 2009).

Electroencephalography (EEG)

Several EEG metrics are used as measures of task engagement (Léger et al., 2014). The most common task engagement metric is the Engagement Index, corresponding to the ratio between beta and the addition of alpha and theta wave power (Pope et al., 1995).

$$Engagement\ Index = \beta/(\alpha + \theta) \tag{1}$$

Additionally, since the beta signal power is associated with a state of alertness and cognitive engagement and the alpha signal power with a state of relaxation, the ratio of beta to alpha is used to reflect arousal levels (Eldenfria and Al-Samarraie, 2019). P3 event-related amplitudes are also often used to measure task engagement because of their close link with motivation and attention (Murphy et al., 2011; Hopstaken et al., 2015).

Heart rate variability (HRV)

HRV is defined as the variation of time intervals between consecutive heartbeats (Mccraty and Shaffer, 2015) and is mainly used as a measure of the activation of the autonomous nervous system (ANS) (Shaffer et al., 2014). Many metrics can be derived from HRV, but the most commonly used are the power of the high-frequency band of HRV (0.15–0.4 Hz) (HF-HRV), the power of the low-frequency band of HRV (0.04–0.15 Hz) (LF-HRV), the ratio of LF-to-HF power, the standard deviation of normal sinus beats (SDNN) and the root mean square of successive RR interval differences (RMSSD). More details on all HRV metrics can be found in (Shaffer and Ginsberg, 2017). To accurately interpret the various measures of Heart Rate Variability (HRV), a brief overview of the Autonomic Nervous System (ANS) is essential. The ANS is governed by two primary components: the parasympathetic and sympathetic systems. The parasympathetic system orchestrates the body’s “rest and digest” responses, promoting relaxation and energy conservation. Conversely, the sympathetic system triggers the “fight or flight” responses, preparing the body for action and mobilizing energy resources. Higher activation of the parasympathetic system is usually associated with better cognitive performance (Hansen et al., 2003) and a better capacity for cognitive engagement (Williams et al., 2016).

HF-HRV reflects parasympathetic activation, where higher HF-HRV is associated with greater activation of the parasympathetic system (Shaffer and Ginsberg, 2017). Since HF-HRV reflects parasympathetic activation, higher HF-HRV is associated with a higher capacity for cognitive engagement.

There is a certain ambiguity concerning the mechanisms underlying the LF-HRV. It may be produced by the sympathetic nervous system, parasympathetic nervous system, or baroreceptors (Shaffer and Ginsberg, 2017). Because of this ambiguity, there is no apparent interpretation of the LF-HRV in the literature. However, because of the potential link between LF-HRV and the sympathetic nervous system, the ratio of LF/HF has been used to reflect the ratio of parasympathetic to sympathetic activation (Shaffer and Ginsberg, 2017). Since parasympathetic activation is linked to better cognitive performance (Hansen et al., 2003; Williams et al., 2016), lower values of LF/HF can be associated with higher capacity for task engagement.

RMSSD reflects the beat-to-beat variance in heart rate and is used to assess short-term heart rate variability. For ultra-short recordings of HRV (under 5 minutes), the RMSSD is correlated with HF-HRV and is usually the primary time domain metric used to estimate the vagally-mediated changes reflected in HRV (Shaffer et al., 2014). Higher RMSSD is typically associated with higher parasympathetic activation, leading to a better cognitive engagement capacity. RMSSD has also been shown to decrease with task difficulty (Hajra and Law, 2020).

SDNN represents the overall variability in heart rate and is usually used to assess global heart rate variability in longer-term HRV measurements. Higher overall variability is associated with a better capacity for cognitive engagement.

Functional near-infrared spectroscopy (fNIRS)

Functional near-infrared spectroscopy measures the change in blood oxygenation in the brain’s cortex and is often used in combination with EEG (Karran et al., 2019; Dehais et al., 2018). The interest in using fNIRS with EEG is to leverage the spatial resolution of fNIRS with the temporal resolution of EEG (Karran et al, 2019). Verdière et al. (2018) have shown that fNIRS signals could be used to detect higher or lower task engagement during a piloting task. They also showed that connectivity measures lead to better classification performance than oxygenation measures.

Electrodermal activity (EDA) and respiration rate (RR)

Electrodermal activity reflects the skin's conductivity level and is used to indicate a state of arousal or stress (Léger et al., 2014; Castiblanco Jimenez et al., 2023). As for respiration, respiration rate has been found to have a significant positive relationship with task engagement (Fairclough and Venables, 2005).

Although eye-tracking and EEG methods are well-established in the literature for assessing task engagement (Pope et al., 1995; Léger et al., 2014; Vadeboncoeur et al., 2024), their practical application in manufacturing faces significant challenges, primarily due to the movement and dynamic environment in which manufacturing operators must operate. Due to these limitations, it is proposed to use measures of alternative metrics, such as respiration rate and HRV. Despite these metrics being less frequently utilized and explored in the literature on human-machine interaction, they are more easily applicable in a manufacturing setting due to their low cost, low intrusiveness, and low sensitivity to movement (Kundinger et al., 2020; He et al., 2022).

2.2.3 Proposed Approach

The constraints inherent to the manufacturing sector make Moray and Inagaki's (2000) approach particularly appealing. Their method evaluates monitoring performance by contrasting actual operator performance to an optimal standard. From this perspective, for any specific task, it seems feasible to establish a performance metric by initially recording the responses of an operator in a high-performance scenario and comparing it to a low-performance scenario. Therefore, in the case of operator engagement, a similar approach would be to establish an engagement metric by comparing physiological responses recorded in highly engaging scenarios with those from a minimally engaging scenario, using contrast to construct a reliable measure of engagement for this task (Couture et al., 2024; Yurish, 2024, p.232). To create high and low engagement scenarios, we can use the approach of Verdière et al. (2018), who manipulated the level of automation to create more and less engaging piloting tasks. This approach is consistent with the findings that showed that higher automation can reduce operator engagement (Passalacqua et al., 2024a).

Hence, to maintain optimal engagement levels of manufacturing operators within their dynamic work environments, our proposal involves developing a new adaptive engagement feedback system inspired by the research of Demazure et al. (2021) but tailored to the manufacturing context. Rather than depending on exact real-time engagement metrics and measurements, our system follows a methodology inspired by the work of Moray and Inagaki (2000), leveraging physiological indicators that differentiate between optimal and suboptimal engagement states. A significant advantage of this approach is its adaptability to complex settings like manufacturing, where constraints exist concerning the feasibility of specific physiological measurements, such as eye-tracking and EEG.

2.3 Objectives

To guide the design process, we established three research objectives: (i) Identify the most appropriate physiological tools for measuring task engagement in a manufacturing context; (ii) Identify and characterize the physiological differences between “high” and “low” engagement manufacturing scenarios; and (iii) Develop an interface that intuitively translates the identified physiological markers into a color gradient, offering immediate feedback on engagement levels. While developing our system, we encountered two significant design challenges that needed careful consideration. The first challenge concerned the optimal display modality for the color gradient, which is crucial for providing clear and understandable feedback on engagement levels. The second challenge involved devising an engagement index scaling method that accurately reflects engagement levels, ensuring that the system's feedback is both intuitive and effective. To address these challenges, we introduced two additional objectives: (iv) Determine the most effective visual representation of engagement through a comparative analysis of a continuous color gradient with 100 shades versus a discrete color gradient with three distinct colors, and (v) Identify the optimal method for scaling the engagement index that accurately reflects perceived engagement, facilitating easier interpretation of physiological markers of engagement by users. With the addition of these two objectives, we were able to make informed design choices that significantly enhanced the usability and effectiveness of our system.

2.4 Methods

We used a design science methodology to develop a task engagement feedback system involving a six-step process that included three studies (see **Table 2.1**). We first selected non-intrusive physiological tools to measure task engagement in a manufacturing assembly context. Then, we collected physiological, performance, and subjective data during "high" and "low" engagement manufacturing scenarios. We identified the physiological differences between the "high" and "low" engagement scenarios and used these markers to design a task-specific "engagement index". Using this formula, we developed the first version of the feedback system. We then evaluated the best display modality and the best scaling method for our engagement index, which were critical aspects of our feature selection process.

Table 2.1 Methodology Employed to Design the Feedback System

Step	**Step 1**	**Step 2**	**Step 3**	**Step 4**	**Step 5**	**Step 6**
Title	***Select Physiological Tools***	***Collect Data***	***Identify Markers***	***System Design***	***Display Validation***	***Scaling Validation***
Description	Comparative analysis of task engagement measures collected with EEG, fNIRS, ECG, eye-tracking and breathing bands	Study 1: Collection of Physiological Data in Scenarios with Varied Engagement Levels	Identify physiological markers of engagement	-	*Study 2:* Validating multiple display modalities of engagement	*Study 3:* Validating multiple index scaling methods
Experimental design	-	Within-subject	-	-	Within-subject	Between subject
Conditions	-	No automation Automation	-	-	Discrete color gradient (3 shades of color) Continuous color gradient (100 shades between green and red)	Min/Max since the beginning of the task Min/Max of training data Min=25th and Max=75th percentiles since the beginning of the task
Experimental manipulation	-	Manufacturing QandA and assembly tasks using snowshoes	Feature extraction using a logistic regression model Validation with LOOCV	-	Fully automated manufacturing QandA and assembly tasks using images of snowshoes	Fully automated manufacturing QandA and assembly tasks using images of snowshoes
Data	Literature review	Collected physiological data (bpm, breath rate, motion) and perceived work engagement	Task 1 and Task 2 data from step 1	-	10 minutes semi-directed interviews	Three-item questionnaire

		(UWES)				
Participants	-	22 participants	-	-	3 participants	10 participants

Step 1 - Choosing Physiological Tools Suitable for a Manufacturing Environment

A thorough methodological reflection was necessary to select the tools and measurements most suited for a manufacturing environment. Our selection criteria dictated that (i) the tool must be non-intrusive for a manufacturing assembly context, (ii) ensure easy data collection, and (iii) provide reliable measurements. Since EEG and fNIRS require wearing a headset, these technologies were deemed too intrusive and potentially distracting for operators in a manufacturing context. Additionally, these technologies are highly sensitive to movement, which is not ideal in a manufacturing setting where the operator must perform physical work. Moreover, electrodermal activity is typically collected on the palm, which would have restricted operators in their assembly tasks. For this reason, EDA was also dismissed for intrusiveness. Given that manufacturing operators often need to interact with a 3D environment, static eye-tracking devices were ruled out. We conducted a pilot test with Tobii Pro glasses (Tobii Technologies, Danderyd, Sweden) that allow the collection of eye-tracking data for users in movement. However, we concluded that using these glasses would overly complicate data collection due to the low battery life and the lack of available tools for analyzing operators' attention when interacting with a 3D environment. This resulted in a preference for electrocardiography and respiration measurements. The Hexoskin vest (Carré Technologies, Montreal, Canada) was found to be a non-intrusive tool that allowed for the simultaneous measurement of these two parameters, as well as accelerometry data. Moreover, heart rate and respiration rate measurements obtained from the Hexoskin vest show little variation compared to laboratory-grade electrocardiograms and metabolic carts, as evidenced by Cherif et al. (2018). Given its accuracy and non-intrusiveness, the Hexoskin vest was selected for the development of our system.

Step 2 – Collect Data in More and Less Engaging Manufacturing Scenarios

In this step, we collected physiological and perceptual data from participants in more and less engaging manufacturing situations. We recruited 22 participants (age=21.62±3.17; men=14) for a within-subject experiment, in which they twice performed a quality control and assembly task on a simulated assembly line. All participants provided a signed consent in-line with the University ethics committee (project # 2023-5058) and were compensated with 40 euros. The task explained in more detail in (Passalacqua et al., 2024a), required participants to detect errors on partially assembled snowshoes and complete the assembly by fixing the binding to the base at its pivot point (see **Figure 2.2**). In the “less engaging” condition, we automated the participants’ decision-making, equipping them with a fully reliable error detection system that indicated to the operator whether a snowshoe had a defect. In the “more engaging” condition, participants had to manually detect errors before assembling the snowshoes. During each task, a total of 30 snowshoes had to be assembled by the participants, with six being defective. Participants realized the task once with automated support and once without automated support, with condition order being randomly assigned and counterbalanced. During the task, we collected physiological data using a Hexoskin vest, recording heart rate, respiration rate, and acceleration data. We also collected perceived cognitive absorption, vigor, and dedication using the Utrecht Work Engagement Scale (UWES) (Schaufeli et al., 2003), which was collected post-task. Since our study specifically aims to modulate the cognitive aspect of work engagement, the absorption dimension is employed as a subjective measure of task engagement within our study. The raw physiological data from the Hexoskin was pre-processed and synchronized using the COBALT Photobooth software (Léger et al., 2019). The list of physiological and self-reported data collected can be found in **Table 2.2**.

Table 2.2 List of Collected Variables

Type of data	**Measure**	**Description**
Physiological data	Beats per minute	Number of beats per minute
	SDNN	Standard deviation of NN intervals
	LF	Power of the Low-frequency band (0.04-0.15 Hz) (ms^2)
	HF	Power of the High-frequency band (0.15-0.4 Hz) (ms^2)
	LF/HF	Ratio of Low-to-High frequency power
	Breathing Rate	Number of respirations per minute
	Minute Ventilation	Respiratory volume per minute (L/min)
	Cadence	Number of steps per minute
	Motion	Norm of the 3D acceleration vector (G)
Self-reported measures	Absorption score (UWES)	Perceived absorption (cognitive engagement)

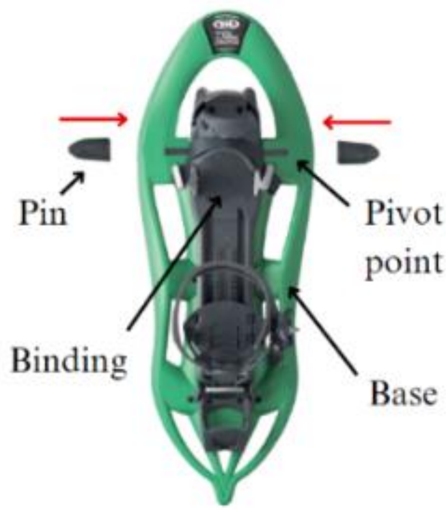

Figure 2.2 - Product Used in the Manufacturing Task

Step 3 – Identify Physiological Markers of Engagement and Create an Engagement Index

In this step, we began by validating our primary assumption that the condition with automation was less engaging than the manual condition. We compared the perceived absorption scores between automated and manual conditions using a one-sided Wilcoxon signed-rank test, which is suitable for evaluating non-parametric paired data. The analysis revealed a statistically significant difference in perceived absorption scores when comparing manual and automated conditions ($p=.0008$), with the automated condition showing lower perceived absorption scores than the manual condition. This result aligns with our primary assumption that the automated condition was less engaging than the automated condition. Based on this result, we then categorized the data, assigning labels of "high" or "low" engagement to arrays of data, depending on the condition experienced by the participant. Data from the automated task was labeled as "low engagement", while data from the manual task was labeled as "high engagement". We then defined a task-specific engagement index (TS-EI) using the three physiological variables with the highest estimated weights in the logistic regression model used to predict the probability of a participant being more engaged in the task (see **Formula 2**). The formula represents a weighted sum, where each coefficient corresponds to the respective variable's estimated power to predict if a participant is in a “high” or “low” state of engagement. The formula is based on 30-second data windows.

$$TS\_EI = (435.7\ Motion_{std}) - (175.4\ Motion_{mean}) + (0.78\ RespirationRate_{std}) \quad (2)$$

Without a testing dataset, we validated Formula 2 using the Leave-One-Out Cross-Validation (LOOCV) on the same dataset. We employed the LOOCV in a logistic regression model to predict if a participant’s engagement during a task was “higher” or “lower”. The results of this test demonstrated an average predictive accuracy of 80.95% on the leave-out samples.

Step 4 – Design the Feedback System

In this step, we developed an initial version of the feedback system. To guide our development process, we established 5 main requirements: (i) The system must collect the user's respiration rate and acceleration data in real-time, (ii) communicate the user's task engagement in real-time using a color gradient, (iii) the displayed color must represent the operator's perceived engagement level, (iv) the system must be easy to interpret, and (v) it should not distract the operator during their task. An overview of the designed system can be found in **Figure 2.3**.

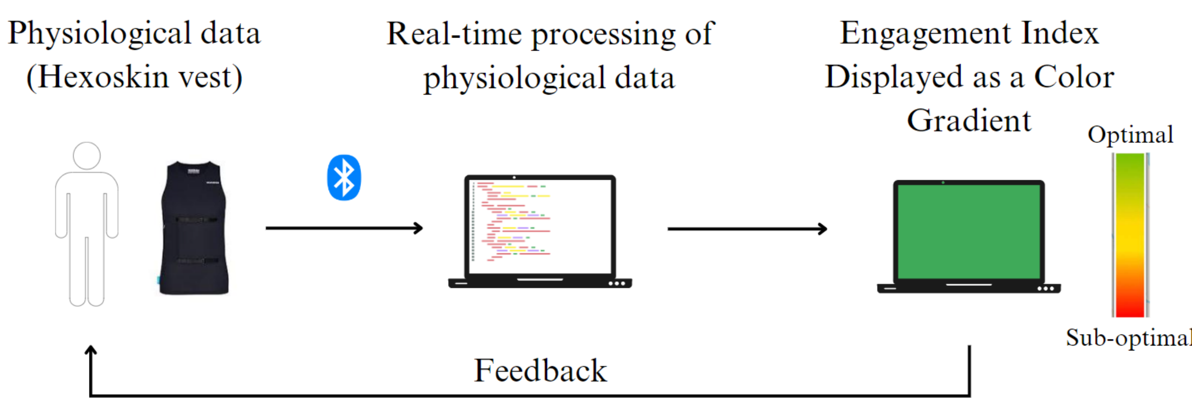

Figure 2.3 - Overview of the Adaptive Feedback System

The system, developed in Python, receives respiration and acceleration data from the Hexoskin vest, which transmits data at a frequency of 1 Hz. Specifically, respiration rate data reflects the average number of respirations per minute based on the last seven breathing cycles, and acceleration data represents the average norm of the 3D acceleration vector over the last second. Our system received data encoded in UTF-8 through a Bluetooth Low Energy (BLE) connection directly established with the Hexoskin vest. It was possible to establish a direct connection using the UUID keys of the respiration and acceleration Bluetooth channels available in Hexoskin's documentation. The system included a Bluetooth reconnection protocol in case of connection failure. Formula 2 was used by the system to calculate the task-specific Engagement Index based on 30-second data windows (or 30 data points, considering that the frequency of transmission of the Hexoskin is 1 Hz). In the first version of the system, the index was normalized using the minimum and maximum index values recorded since the beginning of the session and then scaled as an integer between 0 and 100. Based on the normalized index value, it was possible to select the color to be displayed. The color selection varied according to the display modality, mainly whether the color gradient was discrete (with 3 distinct colors) or continuous (with 100 shades of color). For the continuous gradient, we created a 1x100 matrix with a palette of 100 shades ranging from green to red and used the normalized index value to specify the color code to be fetched from the matrix. For a discrete gradient, only three colors were possible: green for normalized index values above 66%, yellow for values between 33% and 66%, and red for values below 33%. The color codes chosen were then sent via WIFI for display at a frequency of 1Hz. The system's architecture and the specific open-source Python Libraries used in the code are detailed in **Figure 2.4** and **Table 2.3**.

Table 2.3 Python Open-Source Libraries Used by the System

Library	Description	License	Link
Bleak	BLE communication	MIT	https://github.com/hbldh/bleak
Colour	Color code generation	BSD	https://pypi.org/project/colour/

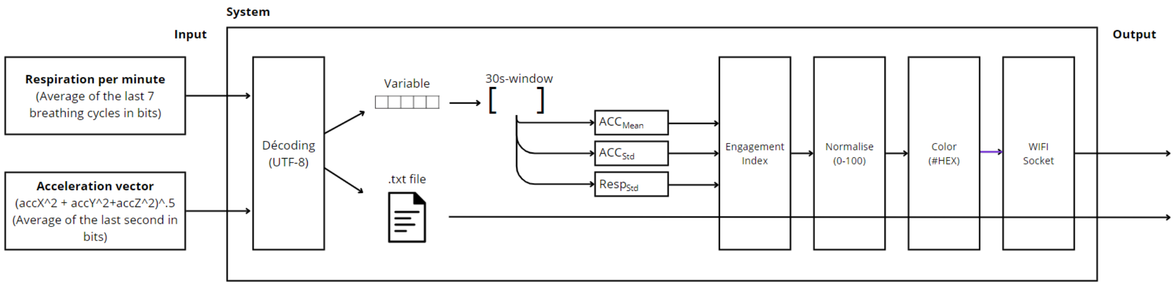

Figure 2.4 - Architecture of the Adaptive Feedback System

Step 5 – Validation of the Display Modality

In this step, we assessed whether representing the index through a continuous color gradient (100 shades) or a discrete color gradient (3 colors) was more effective in conveying participants' engagement levels. We recruited three participants for a within-subjects pilot test. Each participant completed a low-fidelity version of the automated assembly task twice using printed images of snowshoes instead of authentic snowshoes, experiencing the feedback system in both formats. After completing each task, participants underwent a 5-minute semi-directed interview. During this interview, they were asked about the interpretability of the color gradient, the potential distractions caused by the system, and its effectiveness in representing their engagement levels. Positive and negative statements in each category were compiled and analyzed, revealing that the discrete color gradient was more distracting than the continuous color gradient. This led to the decision to retain the continuous color gradient.

Step 6 – Comparative Analysis of Three Scaling Methods

In the sixth step, we aimed to identify the most effective method for scaling the index. We tested three scaling methods: (i) dynamically adjusting the minimum and maximum values based on the minimum and maximum engagement index values recorded since the beginning of the task for this operator, (ii) using the minimum and maximum values of the training dataset, measured with **Formula 3** to exclude outliers, and (iii) dynamically setting the minimum and maximum values respectively to the 25th (Q1) and 75th (Q3) percentile of the engagement index data measured for this operator since the beginning of the task. A visual representation of each method can be found in **Figure 2.5**.

$$MIN/MAX = TS\_EI_{mean} \pm 3 * TS\_EI_{std} \tag{3}$$

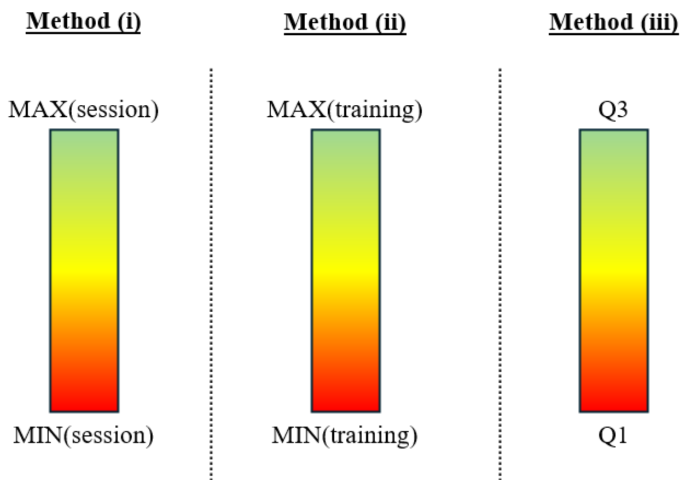

Figure 2.5 - Visual Representation of the Three Index Scaling Methods

We performed a between-subjects experiment with 10 participants who each completed a low-fidelity version of the manufacturing task while receiving feedback from the system in one of its three possible formats. For this low-fidelity version of the manufacturing task, we asked users to identify errors on printed images of snowshoes instead of real snowshoes. After completing the task, participants were asked to rate the color display's representativeness, interpretability, and distractive nature on a scale from 0 to 100. Method (ii) emerged as the most representative of perceived engagement, leading to its selection for the final design. No differences were found in interpretability and distractive nature between the three methods.

2.5 Results

The one-sided Wilcoxon signed-rank test applied in step three demonstrated a statistically significant difference in perceived absorption scores between automated and manual conditions ($p = .0008$). This finding suggests that the distribution of the difference of absorption between automated and manual conditions, is not symmetric around zero, predominantly featuring negative values. This asymmetry suggests that perceived absorption scores are typically lower in the automated condition than in the manual condition, which supports the primary assumption that the automated condition was less engaging than the manual condition. **Table 2.4** and **Figure 2.6** offer an overview of the distribution of reported absorption scores.

Table 2.4 Descriptive Analysis of UWES Absorption Scores Between Manual and Automated Conditions

	Min	Q1	Med	Q3	Max	Mean	Std
Manual	2.67	3.67	4.00	5.25	6.00	4.21	0.99
Automated	1.00	2.67	3.17	4.25	5.33	3.32	1.13

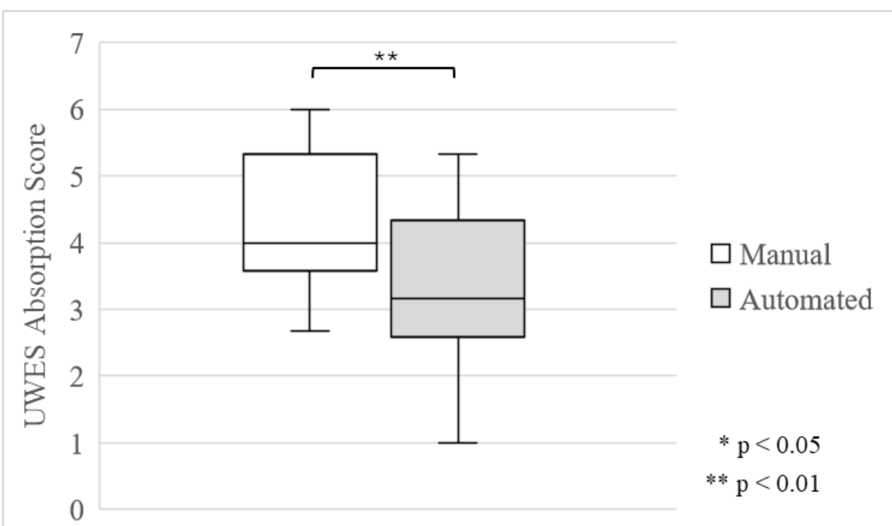

Figure 2.6 - UWES Absorption Scores Distributions Between Manual and Automated Conditions

Using **Formula 2** to predict if a participant was in a “high” or “low” state of engagement in a logistic regression model, we achieved an average of 81.31% accuracy on the training set and 80.95% on the testing set, as confirmed through leave-one-out cross-validation. For step five, where we assessed the display modality, we employed a qualitative labeling technique to categorize interview statements into three themes: effect on perceived engagement, distraction, and representativeness. The number of statements in each category was then compiled (see **Table 2.5**), showing that the discrete color gradient was more distracting (0 positive, six negative statements) than the continuous color gradient (2 positive, 0 negative statements).

Table 2.5 Compilation of Qualitative Statements on Continuous and Discrete Color Gradients

	Perceived effect on engagement		**Distraction**		**Representa-tiveness**	
	(+)	**(-)**	**(+)**	**(-)**	**(+)**	**(-)**
Continuous	5	0	2	0	2	2
Discrete	2	1	0	6	0	3

In step six, the self-reported data from questionnaires revealed that all methods were equally easy to interpret and not distracting. However, the scaling method (ii) utilizing the minimum and maximum values from the training dataset proved to be more representative, with a mean score of 93.33 ±6.24%. This was in contrast to the scaling method (i), which was based on the minimum and maximum values since the beginning of the task (mean= 57.33±12.28%), and method (iii) which was based on percentiles (mean=45.5±14.5%), as illustrated in **Figure 2.7** Based on these analyses, we concluded that the continuous color gradient and scaling method, which utilized the minimum and maximum values of the training dataset, i.e., method (ii), was the preferred option.

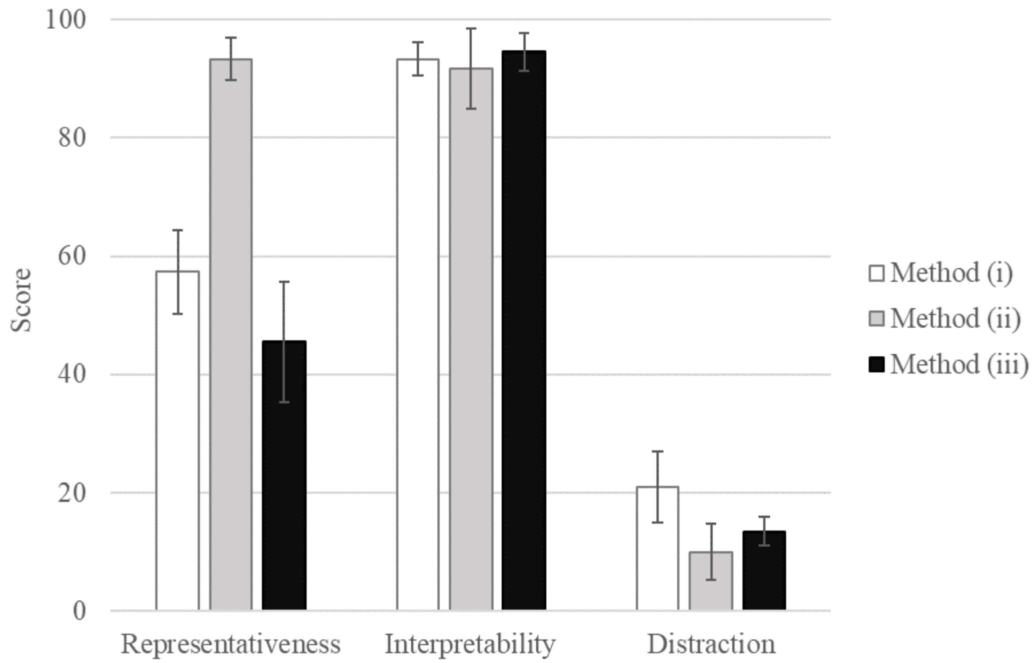

Figure 2.7 - Scaling Method Comparison: Evaluating Representativeness, Interpretability, and Distraction

2.6 Discussion

The objectives of this study were to (i) identify the most suitable physiological tools for measuring task engagement in a manufacturing context, (ii) discern physiological differences between more and less engaging manufacturing situations, (iii) develop an adaptive feedback system that translates these physiological differences into a color gradient for immediate feedback on task engagement, (iv) determine the best mode of displaying engagement between a discrete and a continuous color gradient, and finally (v) find the most representative normalization method for the engagement index as perceived by operators.

For objective (i), we compared eye tracking, EEG, fNIRS, EDA, electrocardiography (ECG), and respiratory rate monitoring tools against three criteria: (a) data collection tools must not distract or disturb the operator during their work, (b) they must allow for easy real-time data collection, and (c) they must provide reliable measurements. EEG, fNIRS, and EDA systems were deemed unsuitable for manufacturing due to their intrusiveness and limitations in dynamic settings. Similarly, static eye-tracking systems failed in 3D environments, and eye-tracking glasses faced battery and analytical challenges. In contrast, ECG and respiratory rate monitoring, conducted via a Hexoskin vest, provided non-intrusive, reliable data collection of engagement metrics, proving effective for manufacturing environments. ECG and respiration metrics are less frequently utilized in the human-machine interaction literature. However, HRV (an ECG metric) has been shown to correlate with well-established engagement metrics such as EEG and eye-tracking, as documented in aviation scenarios by Roy et al. (2016). Additionally, the study by Fairclough and Venables (2005) illustrates that, within their research context, respiration exhibited a stronger correlation with engagement than EEG metrics. While additional validation of respiration as a metric of engagement is required, these findings support the potential utility of ECG metrics and respiration in measuring task engagement.

For objective (ii), we simulated a manufacturing environment and subjected participants to varying engagement levels, using automation to reduce engagement. Participants in the automated condition reported lower absorption scores in the UWES questionnaire, indicating lower perceived cognitive engagement during the automated manufacturing task. This result aligns with previous findings that showed that higher levels of automation can lead to lower task engagement [1].

Based on these findings, we analyzed physiological differences between automated and manual conditions to identify physiological features that could be used to construct a task engagement index. Our observations indicated that participants in the manual condition (condition of higher cognitive engagement) had, on average, lower acceleration means and greater acceleration variability. Without the aid of an error detection tool, participants in the manual condition had to take the time to analyze each product for longer periods than in the automated condition. This contributed to a lower acceleration mean for the manual condition, while the acceleration when fetching a new product increased variability. Considering that a manufacturing operator might not be as focused when moving around as they are when stationary at their worktable, these results suggest the potential use of acceleration mean and acceleration variability as indicators of task engagement. Despite the absence of observable differences in respiration rates across conditions, it was noted that participants engaged in the manual condition (a scenario characterized by greater engagement) exhibited a more consistent respiration rate on average compared to those in the automated condition. This observation is supported by Wientjes (1992), who suggests that rapid shallow breathing is often associated with higher mental workloads and enhanced sustained attention. Consequently, the observed lower variability in respiration rates could be indicative of heightened cognitive effort among participants in the more engaging manual condition. Moreover, Soni and Muniyandi (2019) report a positive correlation between respiration rate variability and heart rate variability, which indicates the potential relationship of this measure with mechanisms underlying task engagement.

For objective (iii), we developed a task-specific engagement index based on the physiological differences explained above. Formula 2 demonstrated good predictive ability on the samples used to create the formula (80.95% predictive capacity).

For objective (iv), displaying the engagement level with a discrete gradient proved less distracting and less representative than using a continuous gradient. This is likely due to the lower sensitivity of the discrete gradient, which affects the operators' sense of control over the system. Additionally, a color oscillation can occur when the measured engagement level approaches a threshold of the discrete gradient, further distracting operators. Conversely, the higher sensitivity of the continuous gradient enhanced the operators' sense of control. It also prevented oscillations between distinct colors, making this method a better alternative for displaying the engagement level.

We compared three methods of normalizing the engagement index. Two of these methods featured dynamic thresholds that were adapted based on data collected since the start of the task, while the other method employed fixed thresholds based on the maximum and minimum values from the training dataset. Results show that the three scaling methods were equally easy to interpret and were not distracting the operators. However, the static threshold method (method ii) was significantly more representative than the two dynamic methods. One possible explanation for this is that the two methods with dynamic thresholds encountered a similar issue where the thresholds diverged as the task progressed, making it increasingly challenging for operators to return to an optimal ("green") engagement level, especially at the end of the task. Therefore, we opted for utilizing the static threshold option for this iteration of the system.

In sum, the adaptive feedback system proposed in this paper utilizes respiration and acceleration data to provide engagement level feedback to manufacturing operators, using a continuous color gradient calibrated using the minimum and maximum engagement values recorded in the training dataset. This system aims to assist manufacturing operators in maintaining optimal engagement levels when interacting with highly automated systems. Providing operators with real-time feedback on their engagement levels ensures they stay informed of their mental state, allowing them to prevent drops in engagement that could adversely impact their performance and, more importantly, safety. The application of this system is particularly relevant in safety-critical manufacturing environments or roles demanding high cognitive engagement, where errors could have significant financial and safety repercussions. A significant benefit of this system is its wide-ranging applicability to various tasks, regardless of their specific characteristics. Additionally, the visual display of engagement can be implemented as an exogenous signal, meaning it does not interfere with the primary task at hand. This versatility underscores the potential of adaptive feedback systems to bolster cognitive engagement during monitoring tasks.

2.7 Conclusion

This study employed a design science methodology to create an adaptive task engagement feedback system designed to help manufacturing operators stay engaged in their evolving workplace. A comparative analysis was utilized to identify the most suitable tools for measuring task engagement in a manufacturing setting, emphasizing the ease of implementation using heart rate variability and respiration rate metrics. A task-specific engagement index was developed using the physiological differences between more and less engaging manufacturing scenarios (acceleration mean, acceleration variability, and respiration variability), achieving an average engagement state prediction accuracy of 80.95% using the leave-one-out cross-validation method in a logistic regression model. We assessed two display modalities and three scaling methods to inform our design. The final design utilized a continuous color gradient calibrated based on the lowest and highest engagement index values recorded in the training set. A subsequent study was conducted to test this advancement on a broader scale, which will be discussed in forthcoming scientific publications.

By offering real-time monitoring and optimization of engagement, this system could help minimize errors and downtime, mitigate safety risks, and promote a healthier work environment.

Thus, it represents a promising approach that could improve both the operational performance and the human experience within manufacturing settings. The theoretical contributions of our work introduce the potential of using measures such as respiration variability and acceleration to infer manufacturing operators' engagement while in motion, as well as the possibility of defining an engagement metric utilizing various physiological differences between optimal and suboptimal scenarios.

It is essential to acknowledge certain limitations inherent in this system. First, our assessment of engagement relied solely on self-reported data. Ideally, employing real-time physiological monitoring tools, like EEG, would have enhanced the validation of the measured engagement levels but would have been more intrusive than the Hexoskin vest we used, potentially distracting operators. Additionally, it should be noted that while the leave-out samples were not employed in training the predictive models, they were utilized in creating Formula 2. As a result, the model's effectiveness for new participants might not be as robust as measured in this study. It is also important to note that the formula used in this system strongly depends on the task and is specifically tailored to the context of our study. This means that Formula 2 may not yield reliable results in different contexts and should not be applied to other scenarios without appropriate modifications and validation. Moreover, using a color gradient can make reading difficult for color-blind users, which affects approximately 8% of the male population. Therefore, in future iterations, it would be important to integrate a color-blindness feature to adjust the displayed colors and improve contrast. Finally, the normalization methods explored in this study did not account for individual physiological differences or natural fatigue occurring during a monitoring task. Regarding individual physiological differences, our study applied a general formula across all participants without differentiation. While effective for establishing a baseline, this approach overlooks the nuances of individual responses and their impact on engagement metrics. Recognizing this limitation, we propose, in further iterations of our research, to refine our engagement threshold criteria by incorporating individual physiological differences into our analysis. This adjustment aligns with the methodology employed by Demazure et al. (2021). As for the fatigue consideration, in our tests, we managed to circumvent the fatigue challenge by conducting short tasks (~15 minutes) where fatigue effects could not realistically take hold. However, employing these methods would result in thresholds that fail to consider fatigue for longer tasks. Therefore, we suggest that future improvements consider the approach outlined by

Demazure et al. (2021) to incorporate fatigue considerations into establishing task engagement thresholds.

References

Allan Cheyne, J., Solman, G. J. F., Carriere, J. S. A., and Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. Cognition, 111(1), 98-113. https://doi.org/https://doi.org/10.1016/j.cognition.2008.12.009

Argyle, E. M., Marinescu, A., Wilson, M. L., Lawson, G., and Sharples, S. (2021). Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments. International Journal of Human-Computer Studies, 145, 102522. https://doi.org/https://doi.org/10.1016/j.ijhcs.2020.102522

Atchley, P., Dressel, J., Jones, T. C., Burson, R. A., and Marshall, D. (2011). Talking and driving: applications of crossmodal action reveal a special role for spatial language. Psychological research, 75, 525-534. https://link.springer.com/content/pdf/10.1007/s00426-011-0342-7.pdf

Bakker, A. B., and Demerouti, E. (2008). Towards a model of work engagement. Career development international, 13(3), 209-223.

Benarroch, E. E. (2009). The locus ceruleus norepinephrine system: functional organization and potential clinical significance. Neurology, 73(20), 1699-1704.

Bernabei, M., and Costantino, F. (2024). Adaptive automation: Status of research and future challenges. Robotics and Computer-Integrated Manufacturing, 88, 102724. https://doi.org/https://doi.org/10.1016/j.rcim.2024.102724

Carter, B. T., and Luke, S. G. (2020). Best practices in eye tracking research. International Journal of Psychophysiology, 155, 49-62. https://doi.org/https://doi.org/10.1016/j.ijpsycho.2020.05.010

Castiblanco Jimenez, I. A., Gomez Acevedo, J. S., Marcolin, F., Vezzetti, E., and Moos, S. (2023). Towards an integrated framework to measure user engagement with interactive or physical products. International Journal on Interactive Design and Manufacturing (IJIDeM), 17(1), 45-67.

Cherif, N. H., Mezghani, N., Gaudreault, N., Ouakrim, Y., Mouzoune, I., and Boulay, P. (2018). Physiological data validation of the hexoskin smart textile.

Couture, L., Passalacqua, M., Joblot, L., Magnani, F., Pellerin, R., and Léger, P.-M. (2024). Enhancing Operator Engagement during AI-assisted Manufacturing Work Using Optimal State Deviation Feedback System.

de Guinea, A. O., Titah, R., and Léger, P.-M. (2013). Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research. Computers in Human Behavior, 29(3), 833-844.

de Guinea, A. O., Titah, R., and Léger, P.-M. (2014). Explicit and implicit antecedents of users' behavioral beliefs in information systems: A neuropsychological investigation. Journal of Management Information Systems, 30(4), 179-210.

Dehais, F., Dupres, A., Di Flumeri, G., Verdiere, K., Borghini, G., Babiloni, F., and Roy, R. (2018). Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. 2018 IEEE international conference on systems, man, and cybernetics (SMC),

Dehais, F., Lafont, A., Roy, R., and Fairclough, S. (2020). A Neuroergonomics Approach to Mental Workload, Engagement and Human Performance [Review]. Frontiers in Neuroscience, 14. https://doi.org/10.3389/fnins.2020.00268

Demazure, T., Karran, A., Léger, P.-M., Labonté-LeMoyne, É., Sénécal, S., Fredette, M., and Babin, G. (2021). Enhancing Sustained Attention. Business and Information Systems Engineering, 63(6), 653-668. https://doi.org/10.1007/s12599-021-00701-3

Eldenfria, A., and Al-Samarraie, H. (2019). Towards an online continuous adaptation mechanism (OCAM) for enhanced engagement: An EEG study. International Journal of Human–Computer Interaction, 35(20), 1960-1974.

Fairclough, S. H., and Venables, L. (2005). Psychophysiological predictors of task engagement and distress. Human Factors in Design, Safety, and Management In D. de Waard, KA Brookhuis, R. van Egmond, and Th. Boersema (Eds.), 349-362.

Feigh, K. M., Dorneich, M. C., and Hayes, C. C. (2012). Toward a Characterization of Adaptive Systems:A Framework for Researchers and System Designers. Human Factors, 54(6), 1008-1024. https://doi.org/10.1177/0018720812443983

Goujon, A., Rosin, F., Magnani, F., Lamouri, S., Pellerin, R., and Joblot, L. (2024). Industry 5.0 use cases development framework. International Journal of Production Research, 1-26.

Gouraud, J., Delorme, A., and Berberian, B. (2018). Out of the Loop, in Your Bubble: Mind Wandering Is Independent From Automation Reliability, but Influences Task Engagement [Original Research]. Frontiers in Human Neuroscience, 12. https://doi.org/10.3389/fnhum.2018.00383

Hajra, S. G., Xi, P., and Law, A. (2020, 11-14 Oct. 2020). A comparison of ECG and EEG metrics for in-flight monitoring of helicopter pilot workload. 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC),

Hallberg, U. E., and Schaufeli, W. B. (2006). “Same same” but different? Can work engagement be discriminated from job involvement and organizational commitment? European psychologist, 11(2), 119-127.

Hansen, A. L., Johnsen, B. H., and Thayer, J. F. (2003). Vagal influence on working memory and attention. International Journal of Psychophysiology, 48(3), 263-274.

He, D., Wang, Z., Khalil, E. B., Donmez, B., Qiao, G., and Kumar, S. (2022). Classification of Driver Cognitive Load: Exploring the Benefits of Fusing Eye-Tracking and Physiological Measures. Transportation Research Record, 2676(10), 670-681. https://doi.org/10.1177/03611981221090937

Hinss, M. F., Brock, A. M., and Roy, R. N. (2022). Cognitive effects of prolonged continuous human-machine interaction: The case for mental state-based adaptive interfaces [Review]. Frontiers in Neuroergonomics, 3. https://doi.org/10.3389/fnrgo.2022.935092

Hopstaken, J. F., van der Linden, D., Bakker, A. B., and Kompier, M. A. J. (2015). The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. Biological Psychology, 110, 100-106. https://doi.org/https://doi.org/10.1016/j.biopsycho.2015.06.013

Karran, A. J., Demazure, T., Leger, P.-M., Labonte-LeMoyne, E., Senecal, S., Fredette, M., and Babin, G. (2019). Toward a Hybrid Passive BCI for the Modulation of Sustained Attention Using EEG and fNIRS [Original Research]. Frontiers in Human Neuroscience, 13. https://doi.org/10.3389/fnhum.2019.00393

Körber, M., Cingel, A., Zimmermann, M., and Bengler, K. (2015). Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. Procedia Manufacturing, 3, 2403-2409. https://doi.org/https://doi.org/10.1016/j.promfg.2015.07.499

Kundinger, T., Sofra, N., and Riener, A. (2020). Assessment of the Potential of Wrist-Worn Wearable Sensors for Driver Drowsiness Detection. Sensors, 20(4), 1029. https://www.mdpi.com/1424-8220/20/4/1029

Léger, P.-M., Courtemanche, F., Fredette, M., and Sénécal, S. (2019). A cloud-based lab management and analytics software for triangulated human-centered research. Information Systems and Neuroscience: NeuroIS Retreat 2018,

Léger, P.-M., Davis, F. D., Cronan, T. P., and Perret, J. (2014). Neurophysiological correlates of cognitive absorption in an enactive training context. Computers in Human Behavior, 34, 273-283. https://doi.org/https://doi.org/10.1016/j.chb.2014.02.011

Léger, P.-M., Karran, A. J., Courtemanche, F., Fredette, M., Tazi, S., Dupuis, M., Hamza, Z., Fernández-Shaw, J., Côté, M., and Del Aguila, L. (2022). Caption and observation based on the algorithm for triangulation (COBALT): Preliminary results from a beta trial. In NeuroIS Retreat (pp. 229-235). Springer.

Lu, Y., Zheng, H., Chand, S., Xia, W., Liu, Z., Xu, X., Wang, L., Qin, Z., and Bao, J. (2022). Outlook on human-centric manufacturing towards Industry 5.0. Journal of Manufacturing Systems, 62, 612-627.

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., Grier, R., and Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry. Emotion, 2(4), 315.

Mazzetti, G., Robledo, E., Vignoli, M., Topa, G., Guglielmi, D., and Schaufeli, W. B. (2021). Work Engagement: A meta-Analysis Using the Job Demands-Resources Model. Psychological Reports, 126(3), 1069-1107. https://doi.org/10.1177/00332941211051988

Mccraty, R., and Shaffer, F. (2015). Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-regulatory Capacity, and Health Risk. Global Advances in Health and Medicine, 4(1), 46-61. https://doi.org/10.7453/gahmj.2014.073

Moray, N., and Inagaki, T. (2000). Attention and complacency. Theoretical Issues in Ergonomics Science, 1(4), 354-365.

Murphy, P. R., Robertson, I. H., Balsters, J. H., and O'Connell R, G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. Psychophysiology, 48(11), 1532-1543. https://doi.org/10.1111/j.1469-8986.2011.01226.x

Naujoks, F., Höfling, S., Purucker, C., and Zeeb, K. (2018). From partial and high automation to manual driving: Relationship between non-driving related tasks, drowsiness and take-over performance. Accident Analysis and Prevention, 121, 28-42. https://doi.org/https://doi.org/10.1016/j.aap.2018.08.018

Parasuraman, R. (2000). Designing automation for human use: empirical studies and quantitative models. Ergonomics, 43(7), 931-951. https://doi.org/10.1080/001401300409125

Parasuraman, R., Molloy, R., and Singh, I. L. (1993). Performance Consequences of Automation-Induced 'Complacency'. The International Journal of Aviation Psychology, 3(1), 1-23. https://doi.org/10.1207/s15327108ijap0301_1

Passalacqua, M., Cabour, G., Pellerin, R., Léger, P.-M., and Doyon-Poulin, P. (2024b). Human-centered AI for industry 5.0 (HUMAI5. 0): Design framework and case studies. In Human-centered AI (pp. 260-274). Chapman and Hall/CRC.

Passalacqua, M., Léger, P.-M., Nacke, L. E., Fredette, M., Labonté-Lemoyne, É., Lin, X., Caprioli, T., and Sénécal, S. (2020). Playing in the backstore: interface gamification increases warehousing workforce engagement. Industrial Management and Data Systems, 120(7), 1309-1330.

Passalacqua, M., Pellerin, R., Yahia, E., Magnani, F., Rosin, F., Joblot, L., and Léger, P.-M. (2024a). Practice with less AI makes perfect: partially automated AI during training leads to better worker motivation, engagement, and skill acquisition. International Journal of Human–Computer Interaction, 1-21.

Pooladvand, S., and Hasanzadeh, S. (2023). Impacts of Stress on Workers' Risk-Taking Behaviors: Cognitive Tunneling and Impaired Selective Attention. Journal of Construction Engineering and Management, 149(8), 04023060. https://doi.org/doi:10.1061/JCEMD4.COENG-13339

Pope, A. T., Bogart, E. H., and Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. Biological Psychology, 40(1-2), 187-195. https://www.sciencedirect.com/science/article/pii/0301051195051163?via%3Dihub

Riedl, R., Fischer, T., Léger, P.-M., and Davis, F. D. (2020). A decade of NeuroIS research: progress, challenges, and future directions. ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 51(3), 13-54.

Rosin, F., Forget, P., Lamouri, S., and Pellerin, R. (2021). Impact of Industry 4.0 on decision-making in an operational context. Advances in Production Engineering and Management, 16(4).

Rosin, F., Forget, P., Lamouri, S., and Pellerin, R. (2022). Enhancing the decision-making process through industry 4.0 technologies. Sustainability, 14(1), 461.

Roy, R. N., Bovo, A., Gateau, T., Dehais, F., and Carvalho Chanel, C. P. (2016). Operator Engagement During Prolonged Simulated UAV Operation. IFAC-PapersOnLine, 49(32), 171-176. https://doi.org/https://doi.org/10.1016/j.ifacol.2016.12.209

Saks, A. M. (2006). Antecedents and consequences of employee engagement. Journal of managerial psychology, 21(7), 600-619.

Scerbo, M. (2007). Adaptive automation. Neuroergonomics: The brain at work, 239252.

Schaufeli, W. B., Bakker, A. B., and Salanova, M. (2003). Utrecht work engagement scale-9. Educational and Psychological Measurement.

Shaffer, F., and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. Frontiers in public health, 258.

Shaffer, F., McCraty, R., and Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. Frontiers in Psychology, 5, 1040. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4179748/pdf/fpsyg-05-01040.pdf

Smallwood, J., and Schooler, J. W. (2006). The restless mind. Psychological Bulletin, 132(6), 946-958. https://doi.org/10.1037/0033-2909.132.6.946

Soni, R., and Muniyandi, M. (2019). Breath Rate Variability: A Novel Measure to Study the Meditation Effects. International Journal of Yoga, 12(1), 45-54. https://doi.org/10.4103/ijoy.IJOY_27_17

Thomson, D. R., Besner, D., and Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. Perspectives on Psychological Science, 10(1), 82-96.

Vadeboncoeur, D., Pellerin, R., and Danjou, C. (2024). Assessing the influence of human factors on overall labor effectiveness in manufacturing: a comprehensive literature review. Automation, Robotics and Communications for Industry 4.0/5.0, 135.

Vasseur, A., Passalacqua, M., Sénécal, S., and Léger, P.-M. (2023). The Use of Eye-tracking in Information Systems Research: A Literature Review of the Last Decade. AIS Transactions on Human-Computer Interaction, 15(3), 292-321.

Verdière, K. J., Roy, R. N., and Dehais, F. (2018). Detecting Pilot's Engagement Using fNIRS Connectivity Features in an Automated vs. Manual Landing Scenario [Original Research]. Frontiers in Human Neuroscience, 12. https://doi.org/10.3389/fnhum.2018.00006

Wientjes, C. J. E. (1992). Respiration in psychophysiology: methods and applications. Biological Psychology, 34(2), 179-203. https://doi.org/https://doi.org/10.1016/0301-0511(92)90015-M

Williams, D. P., Thayer, J. F., and Koenig, J. (2016). Resting cardiac vagal tone predicts intraindividual reaction time variability during an attention task in a sample of young and healthy adults. Psychophysiology, 53(12), 1843-1851.

Yurish, S. (2024). Proceedings of the 4th IFSA Winter Conference on Automation, Robotics and Communications for Industry 4.0/5.0 (ARCI 2024). https://doi.org/10.13140/RG.2.2.20923.1872

Chapitre 3

Enhancing Worker Engagement and Resilience in AI-Assisted Manufacturing: The Role of AR and Engagement Feedback Systems*

Loïc Couture[1], Mario Passlacqua[2], Laurent Joblot[3], Florian Magnani[4], Robert Pellerin[2], and Pierre-Majorique Léger[1]

[1] HEC Montréal, Tech3Lab, 3000 Chemin de la Côte-Sainte-Catherine, H3T 2A7, Canada

[2] Polytechnique Montréal, 2500 Chemin de Polytechnique, H3T 1J4, Canada

[3] Arts et Métiers Institute of Technology, LISPEN / HESAM University / UBFC, France

[4] Université Aix-Marseille, CERGAM, Centrale Méditerranée, 38, rue Frédéric Joliot-Curie, 13013, France

Abstract: In smart manufacturing, keeping workers attentive and engaged is crucial. It is especially important that when automation is interrupted, workers can seamlessly take over the process, maintaining a calm state and sustaining their productivity. However, recent research has shown that integrating artificial intelligence (AI) in manufacturing can sometimes lead to decreased worker engagement, limiting their ability to identify and resolve automation issues effectively. We conducted an experiment to explore the possibility of using cognitive countermeasures to engage workers in AI-supported tasks. We tested two different cognitive countermeasures: augmented reality (AR) and a real-time engagement-level feedback system (RTELFS). The hypothesis is that (i) countermeasures would enhance engagement, motivation, and performance during automated work and that (ii) when automation is removed, workers that used the countermeasures would demonstrate greater resilience, maintaining their engagement, motivation, and performance better than those in the control group. First, we found that the countermeasures helped develop more resilient skills. Workers using the countermeasures showed a smaller reduction in precision when automation was removed. Second, contrary to our expectations, the countermeasures did not impact motivation, cognitive engagement and emotional engagement. Both the control and experimental groups experienced similar levels of motivation,

* Cet article a été préparé pour soumission au journal *International Journal of Production Research (IJPR)*

cognitive engagement and emotional engagement before and when automation was removed. However, countermeasures seemed to positively impact behavioral engagement during the AI-assisted task. Participants using the RTELFS showed increased physical involvement in the AI-assisted task. These findings highlight the potential of cognitive countermeasures to mitigate declines in human performance among workers interacting with automation.

Keywords: Cognitive Countermeasure, Augmented Reality, Engagement Feedback, Resilience, Engagement, Motivation, Manufacturing, Artificial Intelligence

3.1 Introduction

In recent years, there has been increasing deployment of artificial intelligence (AI) systems in smart factories, showing significant increases in both labor and system productivity (Gao and Feng, 2023; Raj and Seamans, 2018). The information processing and projection capabilities provided by AI have enabled systems to become increasingly autonomous, allowing them to be used continuously, beyond traditional work hours (Yang et al., 2021). The integration of AI has also demonstrated positive effects on worker productivity (Plathottam et al., 2023; Raj and Seamans, 2018). AI can be used to reduce the workload of operators by taking over certain repetitive or low-value-added tasks (Tortorella et al., 2024) or by enhancing workers' capabilities, such as providing them with real-time instructions (Sahu, Young, and Rai, 2021).

Currently, the main uses of AI in smart factories include defect detection, predictive maintenance tracking, cost and energy management, as well as the development of robots and autonomous driving systems (Nti et al., 2022), but new applications are gradually emerging. For example, Mypati et al. (2023) recently proposed several AI applications in the areas of casting, forming, and finishing in foundries. Manikandan et al. (2023) highlighted the need to develop AI tools in advanced machining processes and metal welding techniques. Additionally, He et al. (2023) and Mattera, Nele, and Paolella (2024) have recently proposed new ways to use AI in additive manufacturing. With increasingly diverse applications, AI is expected to become increasingly pervasive in smart factories.

When manufacturing operators work with AI systems, it becomes crucial for them to maintain high levels of engagement and attention in their work. Mangler et al. (2021) explain that operators in smart factories often manage a variety of subsystems and interact with multiple levels of automation, which may increase the cognitive load required to perform their tasks (Yamamoto, 2019). This increased mental demand can lead to errors, especially when operators are inattentive (Mangler et al., 2021), which could potentially compromise production quality (Yung et al., 2020) or, in the most severe cases, result in accidents (Naderpour, Nazir, and Lu, 2015). Additionally, when integrating new manufacturing technologies like AI, companies generally expect operators to quickly identify system errors and effectively intervene in the event of automation issues (Endsley and Kiris, 1995). In the event of automation failure, it becomes particularly important that human operators can seamlessly take over the process, maintaining a calm state and sustaining their productivity to mitigate productivity losses (Romero and Stahre, 2021).

However, recent findings show that integrating artificial intelligence (AI) in manufacturing can lead to mixed effects on worker engagement. It can significantly enhance employee engagement when AI is used to reduce the physical and mental efforts of employees in routine tasks, allowing them to focus on tasks that require problem solving, communication and collaboration with colleagues (Tortorella et al., 2024). However, it can lead to decreased operator engagement if AI replaces humans, relegating operators to solely supervisory and passive monitoring roles (Endsley, 2023). One example of this phenomenon is highlighted in a study by Passalacqua et al. (2024). The study found that when operators worked with a fully reliable AI system during manufacturing assembly training, it reduced operator engagement, motivation and ultimately their performance and learning, compared to when training with imperfect AI systems that required human interventions.

One strategy to mitigate this decrease in engagement and performance is to use cognitive countermeasures. Cognitive countermeasures are broadly defined as strategies, techniques or tools that can be used to enhance or maintain cognitive performance (Dehais et al., 2010). They have been used in the past to counterbalance human cognitive bias, such as loss of engagement (Karran et al., 2019) or cognitive tunnelling (Dehais et al., 2010). Although the literature would benefit from having a more specific definition of this concept, in this paper we use the term to refer to strategies and tools that help operators maintain optimal levels of engagement in their work, avoid distractions and mind wandering. In this context, two cognitive countermeasures show great promise for enhancing worker engagement: real-time engagement level feedback systems (RTELFS) and augmented reality (AR).

Real-time engagement-level feedback systems (RTELFS) are systems that measure and assess the level of engagement of workers in real-time using physiological responses (e.g., electroencephalography, heart rate variability) or work performance metrics. They then feedback the information on the level of engagement to the operator, ensuring that operators remain aware of their physiological state of engagement. These systems act as countermeasures by allowing operators to address decreases in engagement in real-time, which has been shown to help maintain optimal cognitive performance during work (Karran et al., 2019). Recently, Couture et al. (2024) have developed a RTELFS that is designed to provide manufacturing operators with real-time information on their level of engagement. However, this system has never been tested in an actual manufacturing context.

A second promising technological solution to enhance operator engagement is augmented reality (AR). AR has the potential to tailor and prioritize the information presented to operators, making it an effective tool for ensuring they remain focused on the most relevant details. For instance, in manufacturing, AR is often used to provide assembly instructions directly within the operator's work environment, minimizing the need to look away and reducing the necessity to divide attention. This approach could be used to alleviate the cognitive load associated with searching for information, increase productivity, and reduce potential distractions (Büschel, Mitschick and Dachselt, 2018). Additionally, AR can deliver information precisely when it is needed, eliminating the need for operators to mentally retain or retrieve details, which can further enhance sustained focus and reduce the risk of mind wandering. AR has demonstrated positive effects on worker engagement in various manufacturing assembly settings (Nguyen and Meixner, 2019; Runji, Lee, and Chu, 2023), but it has not yet been studied as a method to counterbalance potential adverse effects of AI.

Despite the potential of both RTELFS and AR, neither has been studied to mitigate AI-related loss of engagement in smart manufacturing contexts. Therefore, in this study, we aim to explore the potential of using these two countermeasure technologies, both separately and in combination, to enhance the engagement of manufacturing operators in AI-supported tasks. The hypothesis is that if workers are more engaged in their AI-assisted tasks due to the support of countermeasures, they may exhibit higher performance and motivation during these tasks and show greater resilience when required to take over the automated process.

In this work, engagement will refer to a multidimensional concept that encompasses cognitive, emotional, and behavioral dimensions (Mazzetti et al., 2021; Bakker and Demerouti, 2008; Hallberg and Schaufeli, 2006; Saks, 2006; Fredricks et al., 2024). Cognitive engagement refers to the ability to effectively deploy attentional resources toward tasks, leading to optimal focus during work (Pope et al., 1995; Matthews et al., 2002). It involves striking a balance to avoid mind wandering (loss of focus) and cognitive tunneling (an over-focused state that leads to missed signals) (Dehais et al., 2020). Emotional engagement refers to having a positive, fulfilling feeling during work as well as a sense of meaning when realizing the work (Schaufeli, 2013). Finally, behavioral engagement refers the willingness of workers to deploy energy and physical resources during their work (Schaufeli, 2013).

In our experiment, we simulated a snowshoe manufacturing assembly line, drawing inspiration from a visit to an actual snowshoe factory, and introduced a quality assurance and assembly task of 30 snowshoes that had to be completed twice by participants. In the first task, participants were equipped with a highly reliable error detection AI system, along with either AR, RTELFS, a combination of both countermeasures or no countermeasure, depending on their assigned group. In the second task, we removed all technological assistance (i.e., AI and countermeasures) to assess the operators' level of resilience when the automated systems failed. Results show that the countermeasures led to increased resilience in performance when AI assistance was removed, as well as enhanced behavioral engagement in AI-assisted tasks for the RTELFS group.

This paper is divided into seven sections. In Section 2, we provide a review of the literature, offering more detail on countermeasures and their theoretical foundations. In Section 3, we outline the main research objectives and our hypotheses. Section 4 details the methods, including the experimental setup and artifacts used in this study, while Section 5 presents the results obtained for each hypothesis. In Section 6, we discuss the findings, and in Section 7, we provide concluding remarks along with the limitations of the current study.

3.2 Literature review

3.2.1 Cognitive Countermeasures

Cognitive countermeasures are broadly defined as strategies, techniques or tools designed to enhance or maintain cognitive performance during work (Dehais et al., 2010). Countermeasures have been used to refer to different technological support or strategies that helped counterbalance human cognitive bias, such as loss of engagement during work (Demazure et al., 2021; Karran et al., 2019) or cognitive tunneling (Dehais et al., 2010). While the literature would benefit from a more precise definition of this concept, in this paper, we use the term to refer to strategies and tools designed to help operators maintain optimal engagement by reducing the need to divide attention or minimize mind wandering.

One of the first cognitive countermeasures was proposed by Dehais, Causse, and Tremblay (2011), and involved removing information from the dashboard of operators to combat cognitive tunneling, where operators adopt a narrow focus on specific information and risk missing important external signals. Replacing the removed information with signaling cues improved operators' awareness of external signals and enhanced overall decision-making and performance. Cognitive countermeasures have been evaluated in many contexts, ranging from air traffic control (Saint-Lot, Imbert, and Dehais, 2020), autonomous driving (Liu et al., 2024), and even in business process monitoring (Karran et al., 2019). However, to our knowledge, no studies have evaluated the impact of countermeasures in manufacturing contexts.

In the literature, two technologies have shown great potential for enhancing worker engagement and represent significant opportunities to be used as cognitive countermeasures: real-time engagement-level feedback systems (RTELFS) and augmented reality (AR). Due to their potential to enhance worker engagement, these technologies offer promising solutions for mitigating the potential adverse effects of AI on manufacturing operators' engagement. Both systems will be discussed in detail in this section.

<u>Real-Time Engagement-Level Feedback System (RTELFS)</u>

The primary purpose of RTELFS is to help operators stay engaged by enabling them to identify moments of disengagement and prompting an immediate, appropriate response. RTELFS can be seen as cognitive countermeasures because they allow operators to address decreases in engagement in real-time, which has been shown to help maintain optimal cognitive performance during work (Karran et al., 2019). These systems are particularly useful in environments where sustained attention is crucial, such as in manufacturing, healthcare, or control rooms, where maintaining high levels of engagement can directly impact safety, productivity and overall performance.

One notable example of RTELFS was developed by Demazure et al. (2021). Their system used electroencephalography (EEG) to measure the level of cognitive engagement of operators in real-time and update the background color of enterprise resource planning (ERP) software to reflect their corresponding level of engagement. Karran et al. (2019) have explored the potential of using this engagement feedback system to help operators keep optimal levels of sustained attention during a long passive monitoring task of business processes. The results of their study showed that using this countermeasure increased the operator's monitoring performance and EEG wave coherence of operators, which seems to indicate a more sustained level of engagement throughout the task compared to more varied levels of engagement in the no countermeasures groups. Drawing inspiration from these results, Couture et al. (2024) have recently designed an engagement feedback system specifically designed for a manufacturing assembly context. The main advantage of this new system is that it uses physiological metrics such as respiration rate and acceleration that can be more easily collected in manufacturing environments compared to EEG. However, no study has yet been conducted to assess the effectiveness of this countermeasure system for enhancing operator engagement in manufacturing.

To explain why RTELFS could help enhance operator engagement, the theory of operant conditioning seems particularly relevant. Introduced by B.F. Skinner (2019), operant conditioning is a learning process through which behaviors are influenced by reinforcement or punishment. In this framework, behaviors followed by positive outcomes are likely to be repeated, while those followed by negative outcomes are less likely to be repeated. RTELFS can be viewed as a system that provides real-time reinforcement based on the operator's levels of engagement. The feedback serves as a form of operant conditioning, encouraging sustained engagement through positive reinforcement.

Augmented Reality

Another promising solution to mitigate the loss of engagement is augmented reality (AR). AR is a technology that overlays computer-generated information onto real-world environments, enhancing users' perception and interaction with the physical world (Ong, Yuan, and Nee, 2008; Nee et al., 2012). Since AR allows for the manipulation of information accessible to operators, it can be used to ensure that their focus is directed toward the right information, making it a potentially powerful tool to employ as a cognitive countermeasure.

AR has primarily been used in manufacturing settings to provide assembly instructions to operators (Wang et al., 2022; Werrlich et al., 2017). In this context, AR mostly enhances the accessibility and salience of information by overlaying it directly onto the worker’s environment (Romero et al., 2016). This approach has been shown to reduce the cognitive load associated with searching for and interpreting data (Atici-Ulusu, 2021), which could help operators stay focused in their tasks. Additionally, AR can dynamically adapt to the operator's current context and needs, delivering personalized and relevant information precisely when it is needed, further reducing potential distractions, mind wandering or downtime associated with task switching (Romero et al., 2016). To give an example how AR could be used to implement cognitive countermeasures, we could, for instance, use AR to implement the cognitive countermeasure described by Dehais, Causse, and Tremblay (2011), where critical information is temporarily removed by the AR system and replaced with external signals to prevent cognitive tunneling.

AR has shown great potential to enhance operator engagement in manufacturing contexts. Nguyen and Meixner (2019) have shown that gamification using AR in manufacturing assembly training can lead to a general increase in perceived worker engagement. Moreover, Yang et al. (2023) demonstrated that manufacturing assembly training with AR glasses could increase the knowledge retention of operators on a month-term basis. Finally, Lam et al. (2021) showed that the utilization of a smartphone-based AR application increased participants’ knowledge retention compared to a paper-based modality in a product part and disassembly process. Despite its great potential, AR has not yet been evaluated to mitigate AI-related loss of engagement in manufacturing.

The dual-task interference theory (Pashler, 1993) can help explain why AR might lead to better operator engagement. According to this theory, performing two tasks simultaneously can cause interference, particularly when both tasks require similar cognitive resources, leading to decreased performance and potentially lower engagement. The reduction in engagement can be attributed to the increased cognitive load that happens when attention must be divided between multiple tasks (Strayer et al., 2015). This higher cognitive load can lead to overload, which is associated with decreases in worker engagement (Biondi, 2023). From this perspective, if AR can minimize the need to search for information across multiple locations (e.g., consulting a dashboard or referencing a manual), it would reduce dual-task interference, thereby helping to sustain worker engagement.

In sum, cognitive countermeasures are strategies designed to enhance or maintain cognitive performance during work, helping operators prevent sub-optimal cognitive states, such as decreased cognitive engagement. Engagement feedback systems and AR have shown great promise be used as cognitive countermeasures to enhance operator engagement. However, these solutions have not yet been evaluated in smart manufacturing contexts to mitigate AI-induced engagement decrements.

3.2.2 Link Between Engagement, Performance, Motivation and Resilience

Engagement and Performance

Cognitive engagement is a critical predictor of performance, particularly in roles that require high mental involvement, such as supervising automated systems. When operators are cognitively engaged, they are more likely to stay focused and attentive, which enhances their ability to detect and respond to system errors (Endsley and Kiris, 1995). In contrast, a lack of cognitive engagement often leads to distractions or mind-wandering (Dehais et al., 2020), resulting in missed information and errors going unnoticed (Casner and Schooler, 2015). In manufacturing, these undetected errors can have significant consequences. By maintaining high levels of cognitive engagement, operators can adopt better monitoring behaviors, ultimately improving their performance and reducing the risk of missing critical information (Moray and Inagaki, 2000).

Engagement and Motivation

In the literature, employee engagement is often evaluated alongside employee motivation (Mariza, 2016; Pourabdollahian, Taisch, and Kerga, 2012; Latta and Fait, 2016). Employee engagement, in its broad definition, refers to the general willingness of workers to be committed to the organization and their work. Motivation, on the other hand, refers to the various forces that influence employees' levels of commitment, whether it is personal enjoyment of performing the task, external pressure, or rewards. Motivation is commonly regarded as a key driver of worker engagement (Dehais et al., 2020). Consequently, the reverse may also hold true. If countermeasures can enhance an operator's engagement in their work, this increased engagement could, in turn, positively influence their motivation.

The framework used for evaluating motivation in this study is the self-determination theory (SDT). The SDT is a psychological theory concerned with how contextual factors support psychological growth, engagement, and wellness in individuals (Ryan and Deci, 2017). According to this theory, motivation exists on a continuum, with intrinsic motivation at one end, extrinsic motivation in the middle, and amotivation at the other end. Intrinsic motivation refers to the enjoyment derived from simply completing the task and aligns with the individual's interests and values. Extrinsic motivation involves completing the task due to external pressures or incentives. Amotivation, on the other hand, refers to a lack of motivation to complete the task. Van den Broeck et al. (2021) have shown that intrinsic motivation is the strongest predictor of worker well-being and absenteeism. Therefore, this study focuses on intrinsic motivation as the primary metric for evaluating motivation.

<u>Engagement and Resilience</u>

Worker resilience is defined as a person’s capacity to respond to pressure and the demands of daily life (Madni and Jackson, 2009; Peruzzini and Pellicciari, 2017). Workers with higher resilience are generally better equipped to handle stressful situations, such as automation failures. Resilience in smart manufacturing, however, extends beyond human attributes and encompasses resilience in systems as well as the effective integration of human-machine interactions. Recently, Romero and Stahre (2021) introduced the concept of the Resilient Operator 5.0, which combines technical and human dimensions. According to this concept, the Resilient Operator 5.0 is “a smart and skilled operator that uses human creativity, ingenuity, and innovation, empowered by information and technology, to overcome obstacles and create new, frugal solutions.”

For operators to be truly resilient, work engagement appears to be a crucial aspect. Bakker and Demerouti (2008) suggested that engaged employees possess greater adaptability to evolving business conditions mainly because personal attributes such as optimism, self-efficacy, self-esteem, resilience, and an active coping style empower them to manage their work environment efficiently. Moreover, Cooke et al. (2016) have shown that resilience is positively associated with employee engagement in the banking sector. Similarly, Ojo, Fawehinmi, and Yusliza (2021) demonstrated that higher resilience among Malaysian workers during the COVID-19 pandemic was associated with increased work engagement. Therefore, if systems can be implemented to enhance worker engagement, they might also positively impact the resilience of human workers when facing challenging situations such as automation failures.

3.3 Hypothesis Development

The main objective of this study is to explore the possibility of using cognitive countermeasures to engage manufacturing operators in AI-supported tasks. Two specific countermeasures are evaluated: augmented reality (AR) and a real-time engagement-level feedback system (RTELFS) developed by Couture et al. (2024). Since these countermeasures have shown great potential for enhancing worker engagement and that worker engagement has been associated with higher performance, motivation, and resilience among workers, the premise is that if workers are more engaged in their AI-assisted tasks due to the presence of countermeasures, they might be more performant, more motivated and develop greater resilience.

Therefore, without assuming beforehand which method would be superior, we hypothesized that (i) cognitive countermeasures would enhance engagement, motivation, and performance during AI-assisted work, that (ii) when AI-assistance is removed, workers that used the countermeasures would demonstrate greater resilience, maintaining their engagement, motivation, and performance better than those in the control group. This led to the list of hypotheses found below. No specific hypothesis was formulated for the combined effect of the countermeasures, as no theoretical foundation was found to support a prediction. Consequently, the results related to the combination of countermeasures were purely exploratory.

H1: The use of cognitive countermeasures will enhance engagement, motivation, and performance during AI-assisted work.

H1a: During AI-assisted work, the use of cognitive will lead to greater cognitive engagement compared to individuals who did not use countermeasures.

H1b: During AI-assisted work, the use of cognitive countermeasures will lead to greater emotional engagement compared to individuals who did not use countermeasures.

H1c: During AI-assisted work, the use of cognitive countermeasures will lead to greater behavioral engagement compared to individuals who did not use countermeasures.

H1d: During AI-assisted work, the use of cognitive countermeasures will lead to greater motivation compared to individuals who did not use countermeasures.

H1e: During AI-assisted work, the use of cognitive countermeasures will lead to greater performance compared to individuals who did not use countermeasures.

H2: When automation is removed, workers that used the countermeasures would demonstrate greater resilience than those in the control group.

H2a: When automation is removed, workers that used the countermeasures will maintain their cognitive engagement better than those who did not

H2b: When automation is removed, workers that used the countermeasures will maintain their emotional engagement better than those who did not.

H2c: When automation is removed, workers that used the countermeasures will maintain their behavioral engagement better than those who did not.

H2d: When automation is removed, workers that used the countermeasures will maintain their motivation better than those who did not.

H2e: When automation is removed, workers that used the countermeasures will maintain their performance better than those who did not.

3.4 Methods

3.4.1 Participants

A total of 114 students and teachers were recruited for this study (average age = 23.88 ± 7, 21 women). Participants were screened for color blindness, cardiovascular diseases, and skin reactions. All participants provided signed informed consent in accordance with the University Ethics Board (project no. 2023-5427) and were compensated with 40 euros.

3.4.2 Task

For this experiment, we introduced a simple manufacturing assembly task that consisted of inspecting the quality of 30 partially assembled snowshoes and proceeding with assembling only those that met quality standards. There were six possible defects in the products, and participants were trained prior to the task to recognize each defect. The assembly procedure for each snowshoe involved the following steps: (i) click the "Next Snowshoe" button on the dashboard and scan the barcode of the unassembled snowshoe, (ii) inspect the product for defects, (iii) signal any defects via the dashboard or assemble the snowshoe if none were found, and (iv) return the snowshoe to its original position. The assembly manipulation involved attaching the binding to the base of the snowshoe at its pivot (see **Figure 3.1** for an overview of the assembly manipulation).

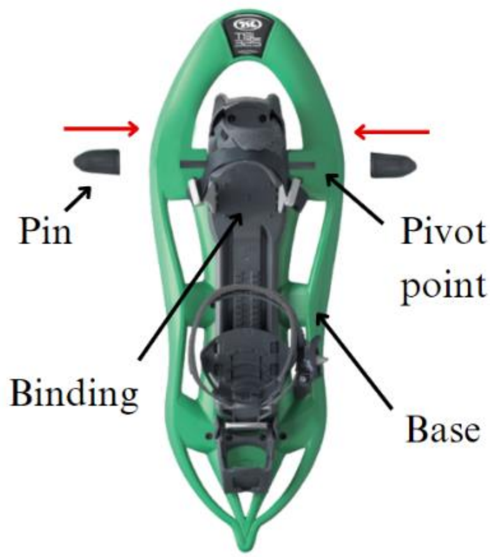

Figure 3.1 - Assembly Manipulation on defect-free snowshoes

3.4.3 Experimental Procedure

The study consisted of one session of 2 hours that was divided into two phases: a training phase and an experiment phase. During the training phase (30 min), participants were provided with a thorough supervised practice of manual fault diagnosis and fault management of six different defaults on snowshoes. The errors included: (i) unattached front binding strap, (ii) inverted front binding, (iii) inverted heel riser, (iv) missing screws, (v) incorrectly mounted spring mechanism for adjusting the binding size, and (vi) back binding strap mounted on the wrong side. The training phase stopped when participants felt confident enough to move to the next step.

During the experiment phase (90 minutes), participants first donned a physiological vest, then completed a demographic questionnaire (age and gender), followed by two consecutive assembly tasks. During the first task, participants were assisted by an AI system that automatically detected errors on the products, thus fully supporting the quality control part of the task. During the second task, we created a scenario where the AI system had failed, requiring operators to detect product errors manually without any assistance from AI. In both tasks, 6 defective products were intentionally introduced into the assembly of 30 snowshoes. A questionnaire was filled by participants after each task.

3.4.4 Experimental Conditions

Participants were randomly assigned to one of three conditions: AR (n=34), RTELFS (n=34) or combined (n=34), which determined the type of technological support they received during the AI-assisted assembly task. The control condition comprised data from 11 participants collected in a previous experiment, using comparable methods, measures, and a similar participant sample. To maintain consistency, we incorporated this existing data into the present study, which accounts for the smaller sample size (n = 11) in the control group.

The control group completed the first task using only the standard AI quality control system without additional technological aids. In the AR condition, participants used the AI quality control system with Augmented Reality. Those in the RTELFS condition utilized the real-time engagement feedback system alongside the traditional AI quality control system. Finally, participants in the combined condition received support from RTELFS and AR alongside the AI quality control system. The second task was identical for all, as it required manual completion without automated support. **Figure 3.2** provides an overview of the experimental design.

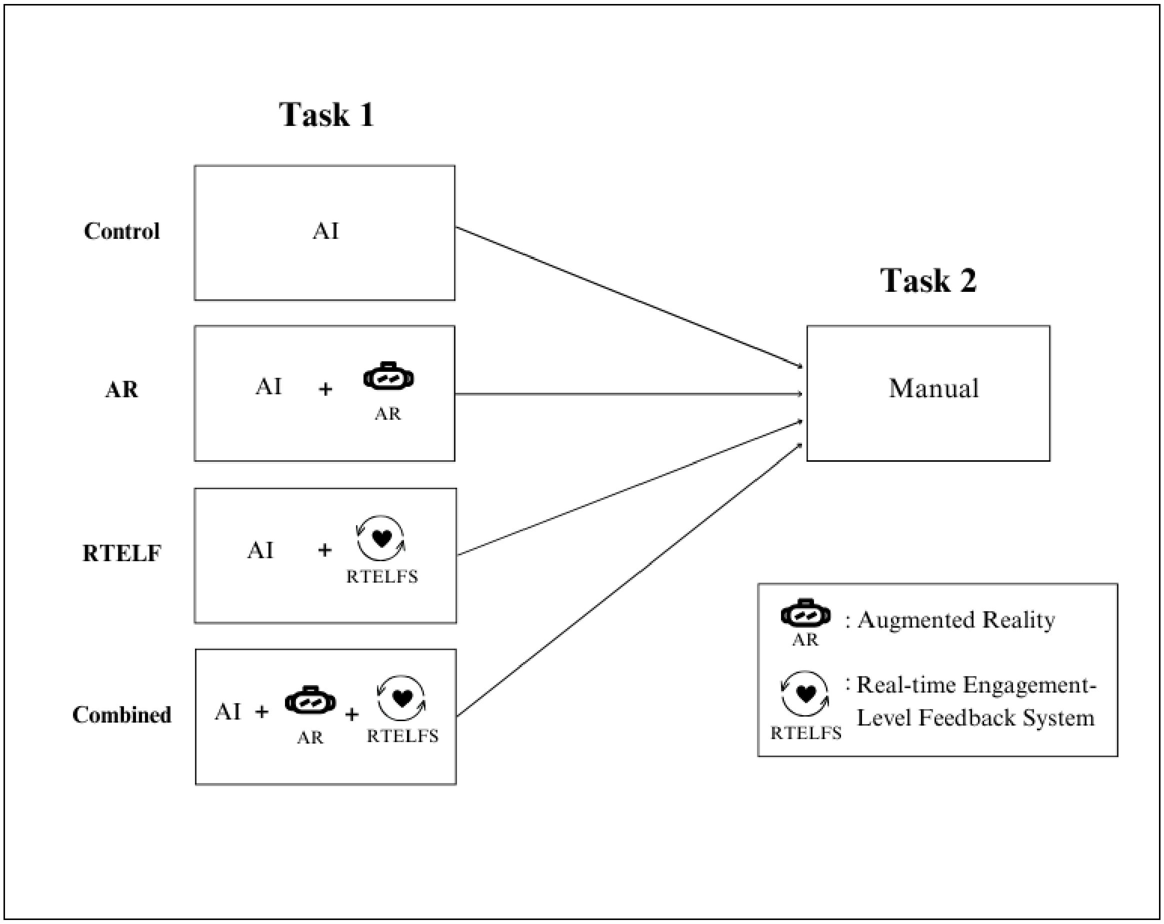

Figure 3.2 - Experimental design

3.4.5 Simulated Assembly Line

Inspired by a visit to an actual snowshoe factory, we developed an assembly line simulation that mirrors the conditions experienced by workers in snowshoe production. The setup included a worktable equipped with a touchscreen dashboard and barcode scanner for interacting with a product management system. This system tracked assembled products and identified defects. Adjacent to the worktable were two racks, each holding 15 snowshoes tagged with barcodes. Participants scanned these barcodes to automatically input data into the product management system. An overview of the workstation is shown in **Figure 3.3**, and the dashboard interface is shown in **Figure 3.4**.

Figure 3.3 - Workstation of the operator

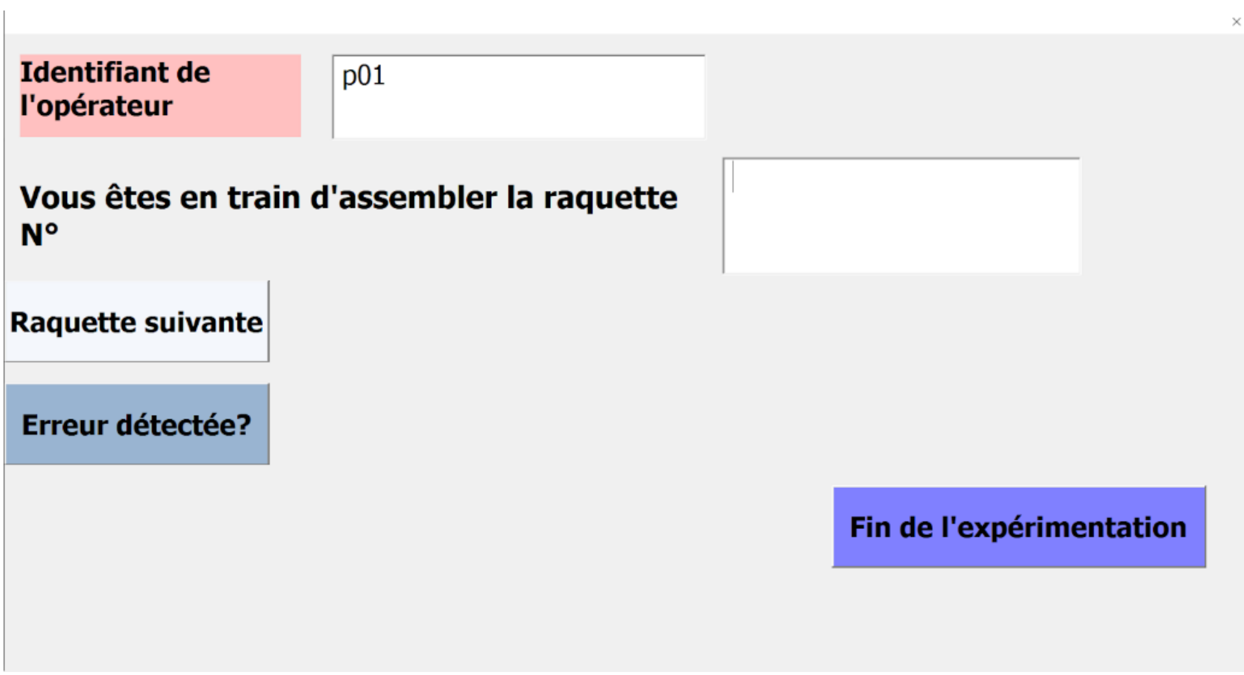

Figure 3.4 - Dashboard Interface of the Product Management System

3.4.6 AI system

To simulate AI assistance, we equipped participants with a fully reliable quality control system designed to automatically detect defects on snowshoes, thus automating the quality control component of the task. Upon scanning a barcode, the AI system displayed diagnostic information directly on the operator's dashboard. This included notifications of detected errors along with detailed specifics, such as a text description of the error, a visual representation using an image, and instructions for the next step (e.g., proceed to assemble the product, scan the error barcode). This system was designed to mimic artificial intelligence capabilities, but it operated under a Wizard of Oz approach (Dahlbäck, Jönsson and Ahrenberg, 1993), employing barcode and QR code analysis to identify pre-planned defects rather than utilizing actual AI technology.

3.4.7 Augmented Reality (AR)

An augmented reality system was specially developed for this experiment to project quality control diagnostic information directly onto the operator's worktable rather than on the traditional operator's dashboard. This system ensured enhanced information accessibility by eliminating the need to turn toward the dashboard. Additionally, it made the information more dynamic, as it interacted directly with the product, using projected arrows that pointed to the identified defect on the product (see **Figure 3.5**). The system utilized a camera and projector, both positioned above the workstation to analyze QR codes affixed to the snowshoes (see **Figure 3.6**) and project diagnostic information onto the worktable. Operators placed a snowshoe within a detection zone, marked by a projected rectangle on the worktable. The camera then scanned the QR codes on the racquet, immediately displaying the diagnostics.

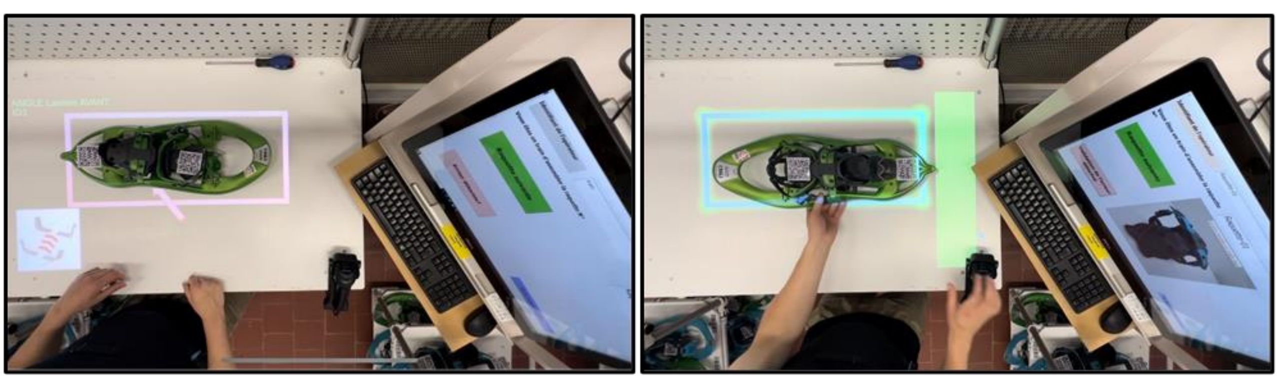

Figure 3.5 - Overview of the Augmented Reality (left) and of the RTELFS (right)

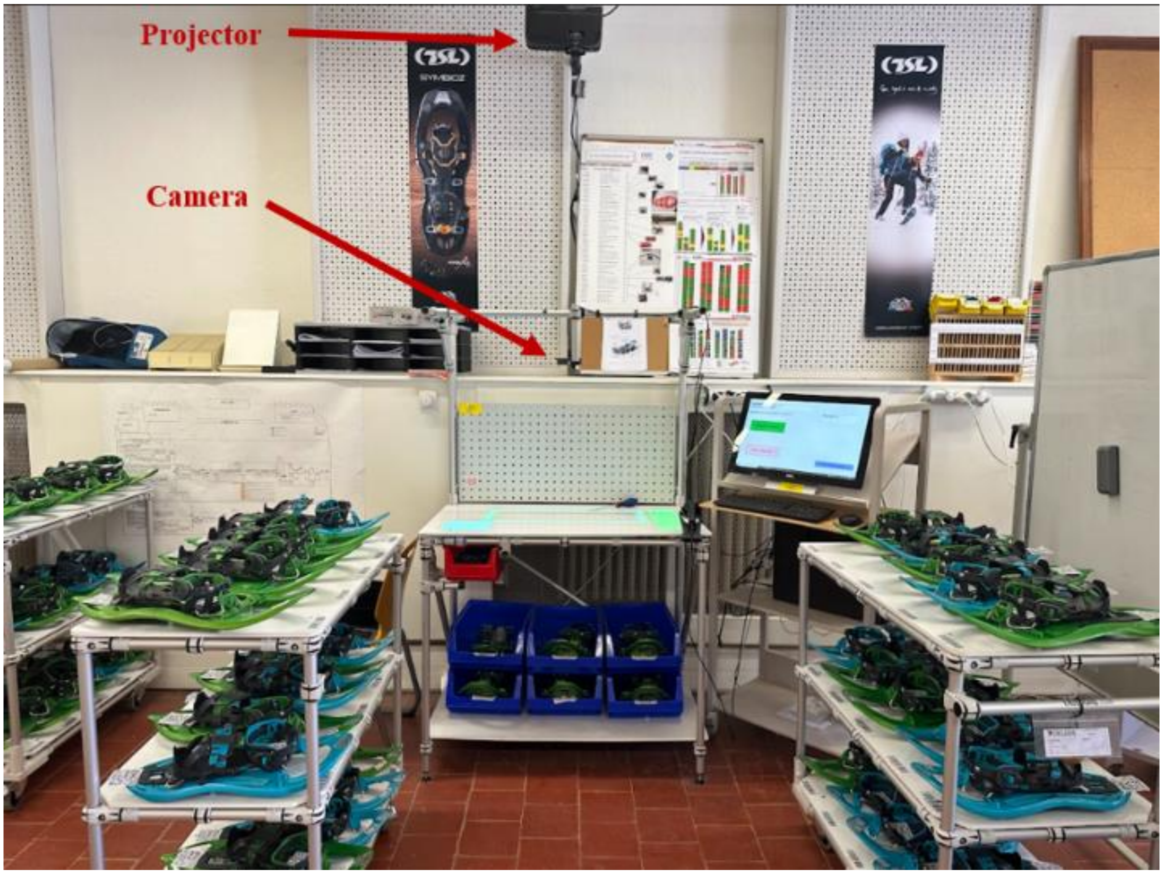

Figure 3.6 - Projected Augmented Reality Setup

3.4.8 Real-Time Engagement-Level Feedback System (RTELFS)

The RTELFS used in this study was a passive engagement level-feedback system developed by Couture et al. (2024). It utilized breathing frequency and acceleration data to intuitively display the operator's level of engagement in real-time through a color gradient. This system was integrated into our augmented reality setup to project the color gradient directly onto the operator's worktable as a large colored rectangle on the side of the table (see **Figure 3.5**). The rectangle's color was updated every second shifting between shades of red and green, reflecting the physiological engagement levels measured by the system. Green indicated high physical and cognitive engagement, while red signaled low engagement. This system was used to maintain participants' awareness of their physiological state of engagement during the task, thereby facilitating necessary adjustments. **Figure 3.7** provides an overview of the RTELF System's functioning.

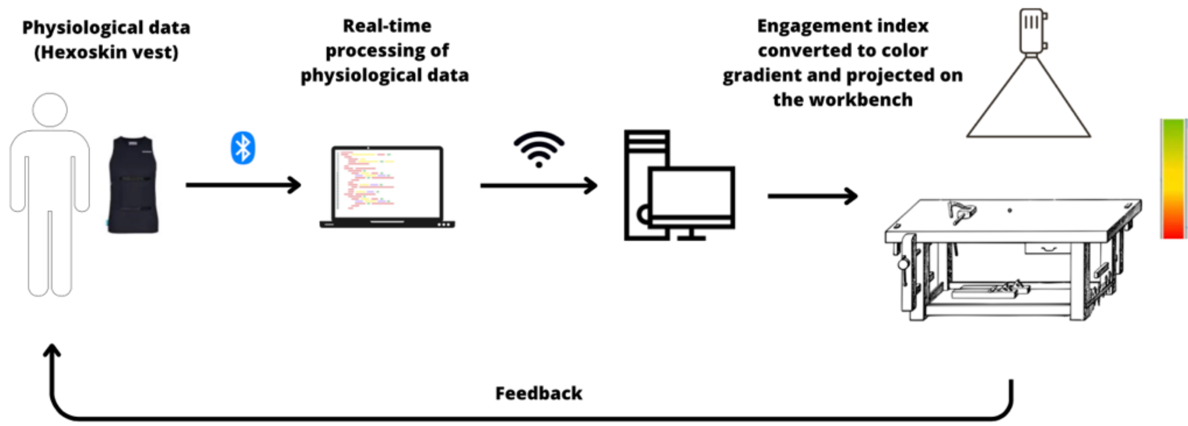

Figure 3.7 - Overview of the Real-Time Engagement-Level Feedback System (RTELFS)

3.4.9 Dependent measures

Eight types of measurements were used to assess the efficiency of the countermeasures. These measures allowed for the evaluation of performance, cognitive, emotional, and behavioral engagement, as well as motivation. The metrics used to evaluate each dependent measure are detailed in the following section.

<u>Performance</u>

Performance was evaluated using the percentage of correct manipulations and completion time for each task. The percentage of correct manipulation was extracted by analyzing the output file of the product management software that compiled all the scans made by the operators. An error of manipulation occurred if a participant reported an error on a valid product or if a participant didn't report an error on a defective product. The percentage of correct manipulation was determined by dividing the number of correctly manipulated snowshoes by the total number of snowshoes (30 per task).

<u>Cognitive engagement</u>

Cognitive engagement was assessed with the absorption subscale of the Utrecht Work Engagement Scale (UWES) and was self-reported in a questionnaire post-task. This subscale is composed of 3 items, rated on a five-point Likert scale. The Absorption dimension of UWES refers to being fully concentrated and deeply engrossed in one's work and is characterized by time passing quickly and difficulties in detaching oneself from work (Schaufeli et al., 2003). In this experiment, we used a French-translated version of the questionnaire, which has been validated by Zecca et al. (2015).

Emotional engagement

Emotional engagement was assessed using valence and arousal sliders. Feldman (1995) defines arousal as a physiologically calm or aroused state (e.g., anxiety, boredom) and valence as a pleasant or unpleasant emotional state (e.g., happiness or sadness)

Behavioral engagement

Behavioral engagement was evaluated using average acceleration and the ratio of low-frequency to high-frequency power of heart rate variability (HRV). Acceleration was used to reflect how physically engaged participants were in the task. The ratio of low-frequency to high-frequency power of HRV is typically used to reflect the ratio of sympathetic to parasympathetic activation of the autonomous nervous system (Shaffer and Ginsberg, 2017). Since the sympathetic nervous system is associated with a state of "fight or flight", while the parasympathetic system is associated with a state of "rest and digest", the ratio can be used to evaluate how much an operator is awake. A higher LF-to-HF ratio would mean more.

Motivation

Motivation was evaluated with the intrinsic motivation subscale of the Empowerment Scale. This subscale is composed of four items, rated on a five-point Likert scale, and was collected in a questionnaire post-task.

Resilience

Resilience was measured by variance between task 1 and task 2 for groups with countermeasures versus Control, for all dependent measures. This included performance, cognitive engagement, emotional engagement, behavioral engagement, and motivation. The lower difference was associated with greater resilience.

3.5 Results

A summary of the mean and standard deviation values recorded for each metric in each group during the AI-assisted task, along with the results of the statistical comparison of means between the countermeasure groups and the control group, is presented in **Table 3.1** Similarly, a summary table containing these results for the task when AI had failed is presented in **Table 3.2**. Additionally, **Table 3.3** provides a summary of the analysis of the differences between task 1 and task 2, as well as the statistical comparison of these differences between the countermeasure groups and the control group.

Table 3.1 Mean and Variance Results in Task 1 (AI-Assisted Task)

Dependent Measure	Metric	Statistical test	Condition	Mean	Std	Diff to Control	p-value Diff to Control
Performance	Manipulation Performance (%)	Mann-Whitney Wilcoxon	Control	99,09	2,16	-	-
			AR	99,14	1,98	0,04	1,000
			BS	98,97	2,69	-0,13	1,000
			Combined	98,82	2,52	-0,27	1,000
	Completion Time (s)	t-test	Control	822	187	-	-
			AR	851	166	29	1,000
			BS	719	132	-103	0,098
			Combined	793	95	-29	1,000
Cognitive Engagement	UWES Absorption Score (/5)	t-test	Control	3,52	0,89	-	-
			AR	3,53	0,87	0,02	1,000
			BS	3,22	0,99	-0,30	1,000
			Combined	3,57	1,04	0,05	1,000
Emotional Engagement	Valence Slider (%)	t-test	Control	54,9	18,5	-	-
			AR	65,2	22,7	10,3	0,491
			BS	64,7	19,2	9,8	0,491
			Combined	59,0	20,4	4,1	1,000
	Arousal Slider (%)	t-test	Control	37,4	17,7	-	-
			AR	40,9	20,0	3,6	1,000
			BS	48,3	25,7	11,0	1,000
			Combined	47,4	24,9	10,0	1,000
Behavioral Engagement	Acceleration Mean (g)	t-test	Control	0,064	0,014	-	-
			AR	0,069	0,012	0,005	0,366
			BS	0,078	0,016	0,015	0,016
			Combined	0,072	0,013	0,008	0,186
	LF/HF ratio (HRV)	t-test	Control	0,691	0,082	-	-
			AR	0,713	0,169	0,022	1,000
			BS	0,617	0,178	-0,074	1,000
			Combined	0,769	0,240	0,078	1,000
Motivation	Empowerment Intrinsic Motivation Score (/5)	t-test	Control	2,75	1,45	-	-
			AR	3,22	1,09	0,47	1,000
			BS	3,20	1,07	0,45	1,000
			Combined	2,95	1,14	0,20	1,000

Table 3.2 Mean and Variance Results in Task 2 (Manual Task)

Dependent Measure	Metric	Statistical Test	Condition	Mean	Std	Diff to Control	p-value Diff to Control
Performance	Manipulation Performance (%)	Mann-Whitney Wilcoxon	Control	90,6	5,5	-	-
			AR	97,3	2,6	6,7	0,004
			BS	97,6	2,9	7,0	0,001
			Combined	96,6	3,6	6,0	0,007
	Completion Time (s)	t-test	Control	1029	295	-	-
			AR	954	162	-76	0,331
			BS	876	153	-154	0,039
			Combined	922	137	-107	0,195
Cognitive Engagement	UWES Absorption Score (/5)	t-test	Control	4,15	1,10	-	-
			AR	3,64	0,96	-0,51	0,491
			BS	3,69	0,94	-0,46	0,491
			Combined	3,66	1,09	-0,50	0,491
Emotional Engagement	Valence Slider (%)	t-test	Control	56,4	24,0	-	-
			AR	62,4	21,7	6,0	1,000
			BS	64,5	17,9	8,1	1,000
			Combined	55,5	21,7	-0,9	1,000
	Arousal Slider (%)	t-test	Control	59,3	21,9	-	-
			AR	66,7	18,1	7,4	1,000
			BS	72,4	15,2	13,2	0,155
			Combined	63,1	21,4	3,9	1,000
Behavioral Engagement	Acceleration Mean (g)	t-test	Control	0,051	0,012	-	-
			AR	0,057	0,011	0,005	0,242
			BS	0,064	0,014	0,013	0,011
			Combined	0,059	0,009	0,008	0,145
	LF/HF ratio (HRV)	t-test	Control	0,73	0,10	-	-
			AR	0,78	0,23	0,05	0,400
			BS	0,66	0,18	-0,07	0,400
			Combined	0,85	0,27	0,12	0,278
Motivation	Empowerment Intrinsic Motivation Score (/5)	t-test	Control	3,27	1,91	-	-
			AR	3,13	1,21	-0,14	1,000
			BS	3,22	1,09	-0,06	1,000
			Combined	2,81	1,04	-0,47	1,000

Table 3.3 Analysis of Variations Between Task 1 and Task 2

Dependent Measure	Metric	Statistical Test (T2-T1)	Statistical Test (Diff with control)	Condition	T2-T1	p-value (T2-T1)	Diff with control	p-value Diff with control
Performance	Manipulation Performance (%)	Mann-Whitney Wilcoxon	t-test	Control	-8,5%	<0,0001	-	-
				AR	-1,9%	0,0021	0,066	<0,0001
				BS	-1,4%	0,0224	0,071	<0,0001
				Combined	-2,3%	0,0028	0,062	<0,0001
	Completion Time (s)	t-test	t-test	Control	207,09	0,0631	-	-
				AR	102,48	0,0258	-104,609	0,190
				BS	156,66	<0,0001	-50,436	1,000
				Combined	128,55	<0,0001	-78,543	0,462
Cognitive Engagement	UWES Absorption Score (/5)	t-test	t-test	Control	0,64	0,1506	-	-
				AR	0,11	0,6592	-0,525	0,109
				BS	0,47	0,0691	-0,165	1,000
				Combined	0,09	0,7522	-0,550	0,077
Emotional Engagement	Valence Slider (%)	t-test	t-test	Control	1,45	0,8753	-	-
				AR	-2,81	0,6436	-4,269	1,000
				BS	-0,24	0,9606	-1,696	1,000
				Combined	-3,58	0,5060	-5,035	1,000
	Arousal Slider (%)	t-test	t-test	Control	21,91	0,0180	-	-
				AR	25,78	<0,0001	3,869	1,000
				BS	24,10	<0,0001	2,194	1,000
				Combined	15,77	0,0097	-6,135	1,000
Behavioral Engagement	Acceleration Mean (g)	t-test	t-test	Control	-0,0124	0,0468	-	-
				AR	-0,0119	<0,0001	0,0005	1,000
				BS	-0,0141	0,0011	-0,0017	1,000
				Combined	-0,0125	<0,0001	-0,0001	1,000
	LF/HF ratio (HRV)	t-test	t-test	Control	0,04	0,3395	-	-
				AR	0,06	0,2456	0,025	0,879
				BS	0,04	0,3630	0,005	1,000
				Combined	0,08	0,2174	0,042	0,260
Motivation	Empowerment Intrinsic Motivation Score (/5)	t-test	t-test	Control	0,52	0,4778	-	-
				AR	-0,09	0,7682	-0,615	0,182
				BS	0,02	0,9517	-0,505	0,355
				Combined	-0,15	0,6023	-0,668	0,115

3.5.1 Performance

Performance During the AI-Assisted Task

Manipulation performances reported during the AI-assisted task were an average of 99.09% for the control group, 99,14% for the AR group, 98.97% for the BF group, and 98.82% for the combined countermeasures group (see **Figure 3.8**). A Mann-Whitney Wilcoxon test was used to compare mean manipulation performance between conditions during the AI-assisted task. The results of this test revealed no statistically significant differences in manipulation performance between the countermeasures group and the control group. For this metric, a non-parametric test was employed because the data was not normally distributed.

No significant differences were found in completion time between the countermeasure groups and the control group, as detailed in **Table 3.1**. However, the BF group completed the AI-assisted task an average of 103 seconds more rapidly than the control group, which represents a 12.5% productivity increase compared to the control group (see **Figure 3.9**). The AR group, on the other hand, took, on average, 29 seconds more than the control group to complete the task, representing a 3.5 % productivity decrease compared to the control group. Finally, the combined countermeasure group completed the task in average 29 seconds quicker than the control group, representing a 3.5% productivity increase compared to the control group.

Resilience of Performance After AI Failure (Manual Task)

After the AI system failed, participants in the control group reported 90.6% manipulation performance, while the AR, BF, and combined group respectively reported manipulation performance of 97.3%, 97.6%, and 96.6% (see **Figure 3.8**). Compared to their performance in the AI-assisted task, this represents an 8.5% decrease in performance for the control group, a 1.9% decrease for the AR group, a 1.4% decrease for the BF group, and a 2.3% decrease for the combined group.

As for completion time, the control group took, on average, 1029 seconds to complete the task, the AR group 954 seconds, the BF group 876 seconds, and the combined countermeasure group 922 seconds (see **Figure 3.9**). Compared to their respective performance in the AI-assisted task, this represents a 207-second increase for the control group, a 102-second increase for the AR group, a 156-second increase for the BF group, and a 128-second increase for the combined group.

Resilience of performance was assessed by comparing these decreases in performance between T1 and T2 for the countermeasure groups against the control group. A t-test with Bonferroni adjustment was used to compare the decrease in manipulation performance between groups. The results of this test revealed that the decrease in manipulation performance in the countermeasure groups was significantly lower than the decrease in performance in the control group ($p<0.0001$ for all conditions). On average, the decrease in manipulation performance was 6.6% lower for the AR group, 7.1% lower for the BF group, and 6.2% lower for the combined group, as compared to the control group. A t-test was also used to compare the loss of productivity (i.e., increased completion time) between the countermeasure groups and the control group. No significant difference was found between the groups.

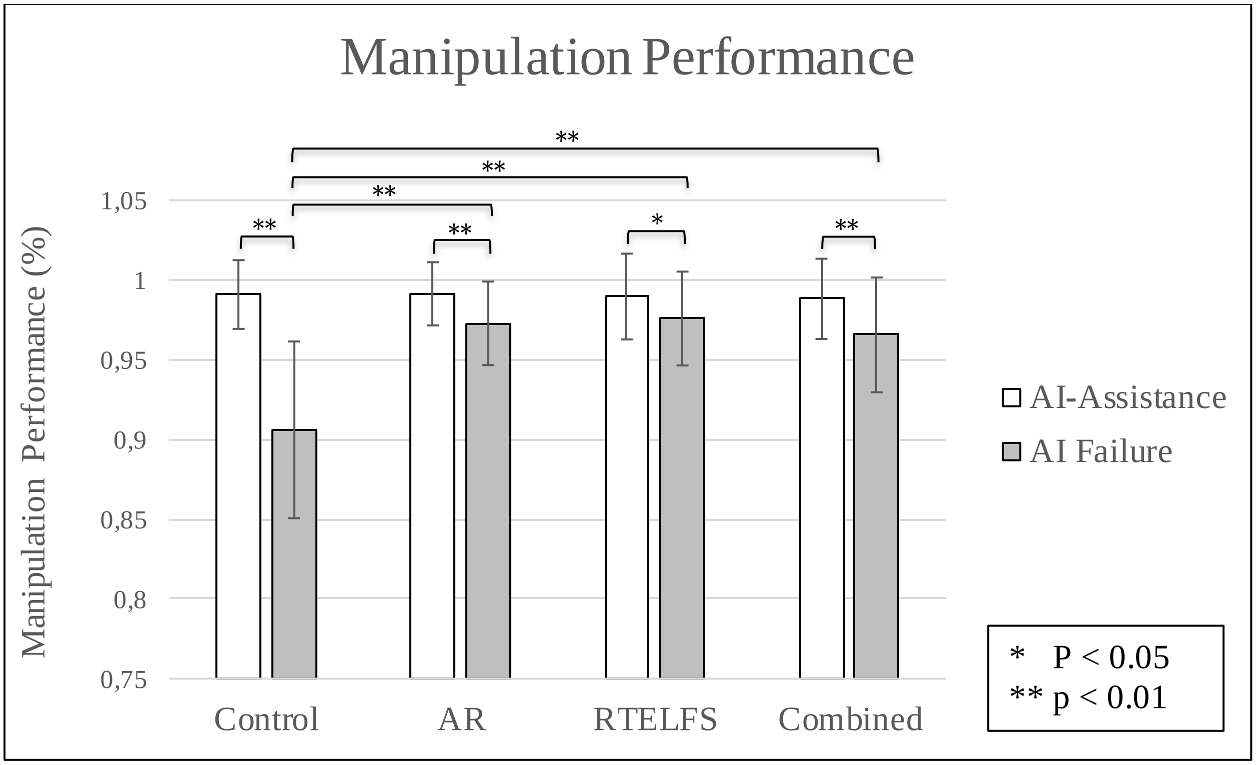

Figure 3.8 – Mean Manipulation Performance Across Groups for Both Tasks

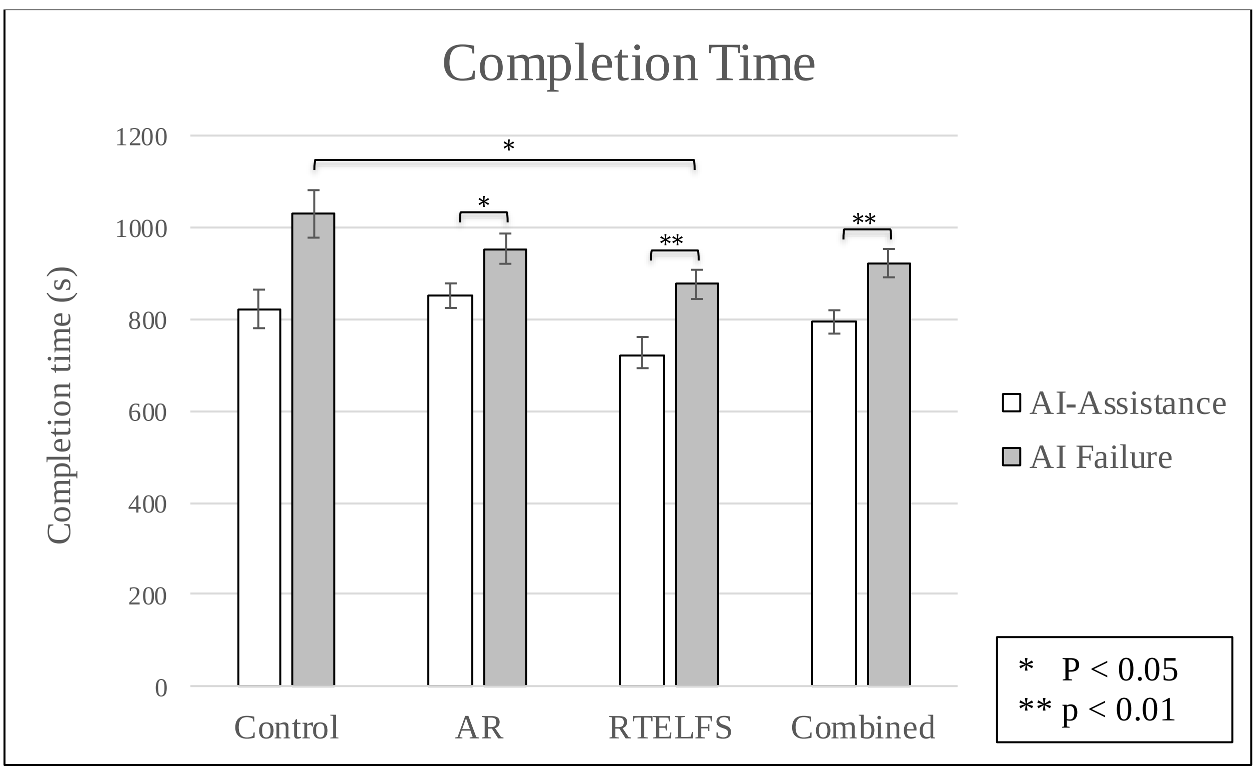

Figure 3.9 – Mean Completion Time Across Groups for Both Tasks

3.5.2 Behavioral Engagement

Behavioral Engagement During the AI-Assisted Task

During the AI-assisted task, the mean acceleration values recorded were 0.064 g for the control group, 0.069 g for the AR group, 0.078 g for the BF group, and 0.072 g for the combined group (see **Figure 3.10**). A t-test with Bonferroni adjustment revealed that the acceleration in the BF group was significantly higher than in the control group (p=.016). No significant differences were found between the other countermeasure groups and the control group.

As for the LF/HF power ratio of HRV, values during the AI-assisted task were 0.69 for the control group, 0.71 for the AR group, 0.62 for the BF group, and 0.77 for the combined group (see **Figure 3.11**). None of the comparisons of LF/HF ratio between the countermeasures groups and the control group showed statistically significant differences.

Resilience of Behavioral Engagement After AI Failure (Manual Task)

When the AI failed, the acceleration mean decreased almost equally between groups. The control group acceleration decreased by an average of 0.012 g, the AR group decreased by 0.012 g, the BS group decreased by 0.014 g, and the combined group by 0.013 g compared to the acceleration in the AI-assisted task (see **Figure 3.10**). A t-test between conditions revealed that this decrease in acceleration was not significantly different between the control group and the countermeasure groups.

As for the ratio of LF to HF power of HRV, we saw small increases in all conditions during the AI failure as compared to the values recorded during the AI-assisted task. The ratio increased by an average of 0.04 in the control group, 0.06 in the AR group, 0.04 in the BS group, and 0.08 in the combined group (see **Figure 3.11**). A t-test with Bonferroni adjustment revealed that this increase in LF to HF ratio was not significatively different in the countermeasure groups when compared to the control group.

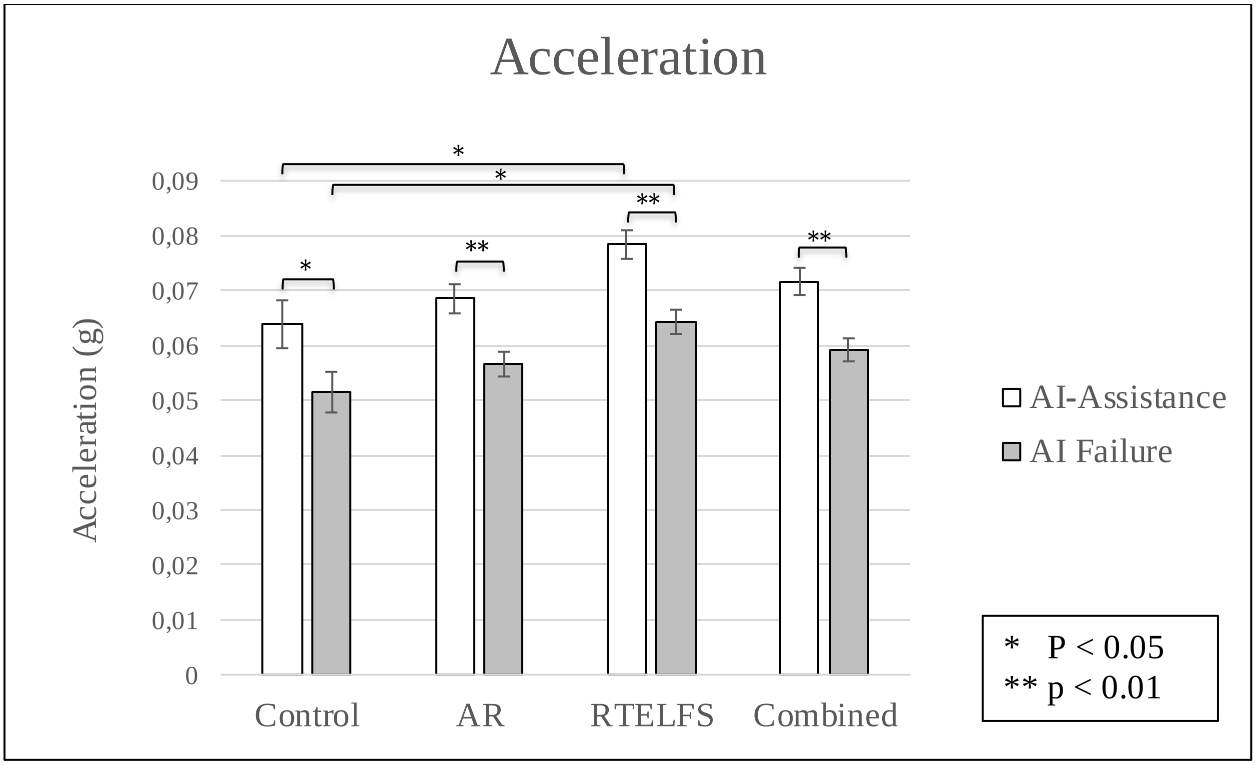

Figure 3.10 – Mean Acceleration Results Across Groups for Both Tasks

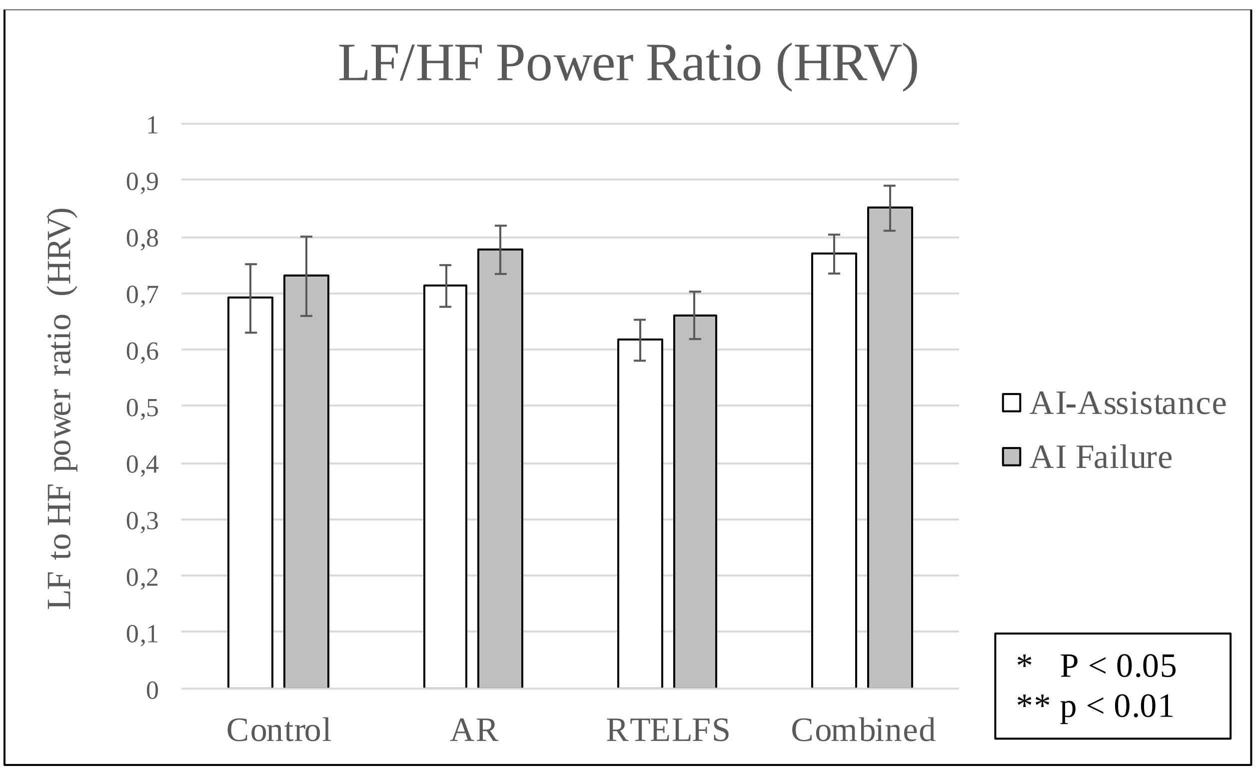

Figure 3.11 – Mean LF/HF Power Ratio of HRV Across Groups for Both Tasks

3.5.3 Emotional Engagement

<u>Emotional Engagement During the AI-assisted task</u>

During the AI-assisted task, the valence slider results were 54.9% in the control group, 65.2% in the AR group, 64.7% in the BS group, and 59.0% in the combined countermeasure group (see **Figure 3.12**). Although the valence was not significantly different in the countermeasures groups compared to the control group, we see that valence was, on average, higher in the countermeasure groups than in the control group. Emotional valence was on average 10.3% higher in the AR group, 9.8% higher in the BS group and 4.1% in the combined group.

For arousal, the control group reported a 37.4% arousal score during the AI-assisted task, while the AR, BS, and combined countermeasure groups respectively reported arousal scores of 40.9%, 48.3%, and 47.4% (see **Figure 3.13**). With these results, we see that arousal was in average higher in the countermeasure groups than in the control group. Arousal was, on average, 3.6% higher in the AR group, 11.0% higher in the BS group, and 10.0% in the combined group as compared to the control group. However, no significant differences between countermeasure groups and the control condition were found.

<u>Resilience of Emotional Engagement After AI Failure (Manual Task)</u>

When the AI failed, valence results slightly increased by 1.45% in the control group, while in the AR, BS, and combined groups, it slightly decreased respectively by 2.81%, 0.24%, and 3.58%, as compared to the AI-assisted task (see **Figure 3.12**). However, this difference in valence between T1 and T2 was not significant for the countermeasures groups compared to the control group.

For arousal, results show a significant increase during AI failure as compared to the AI-assisted task across all conditions. Arousal increased by 21.91% in the control group, 25.78% in the AR group, 24.10% in the BS group, and 15.77% in the combined group (see **Figure 3.13**). However, these increases in arousal were not significantly different in the countermeasures groups compared to the control group.

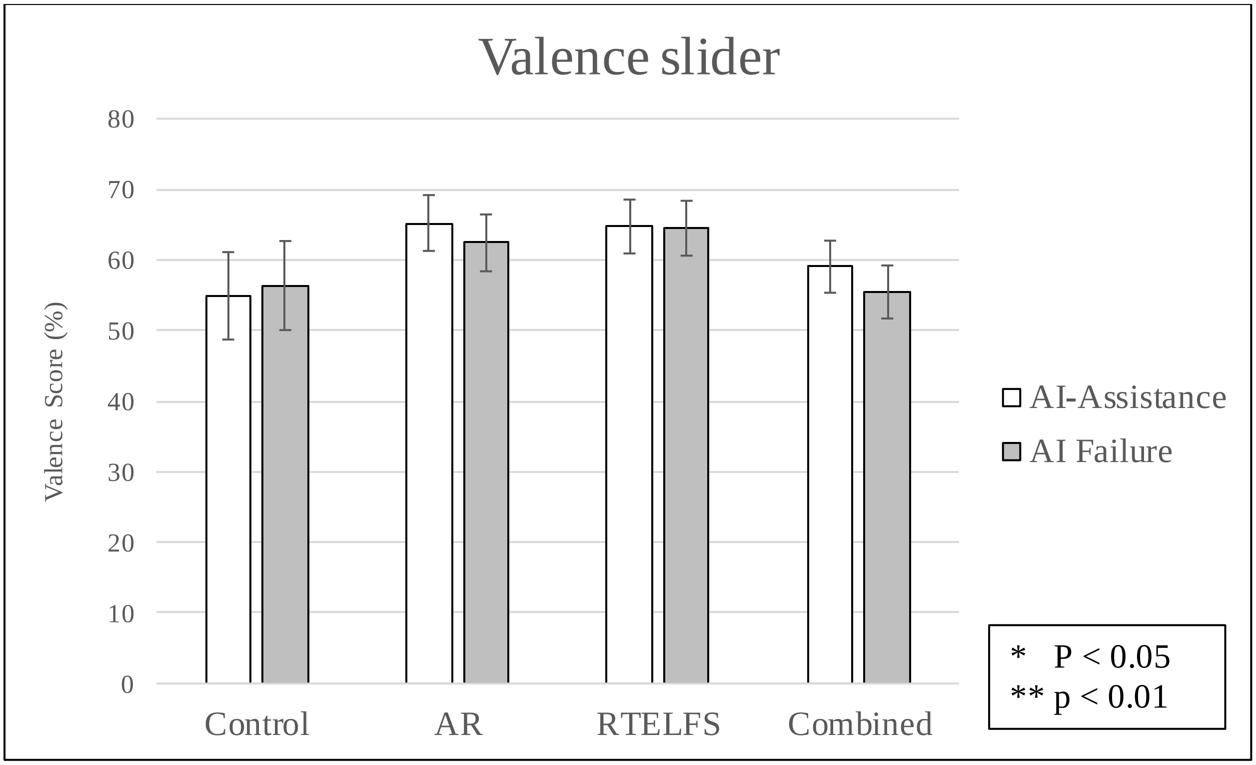

Figure 3.12 – Mean Valence Results Across Groups for Both Tasks

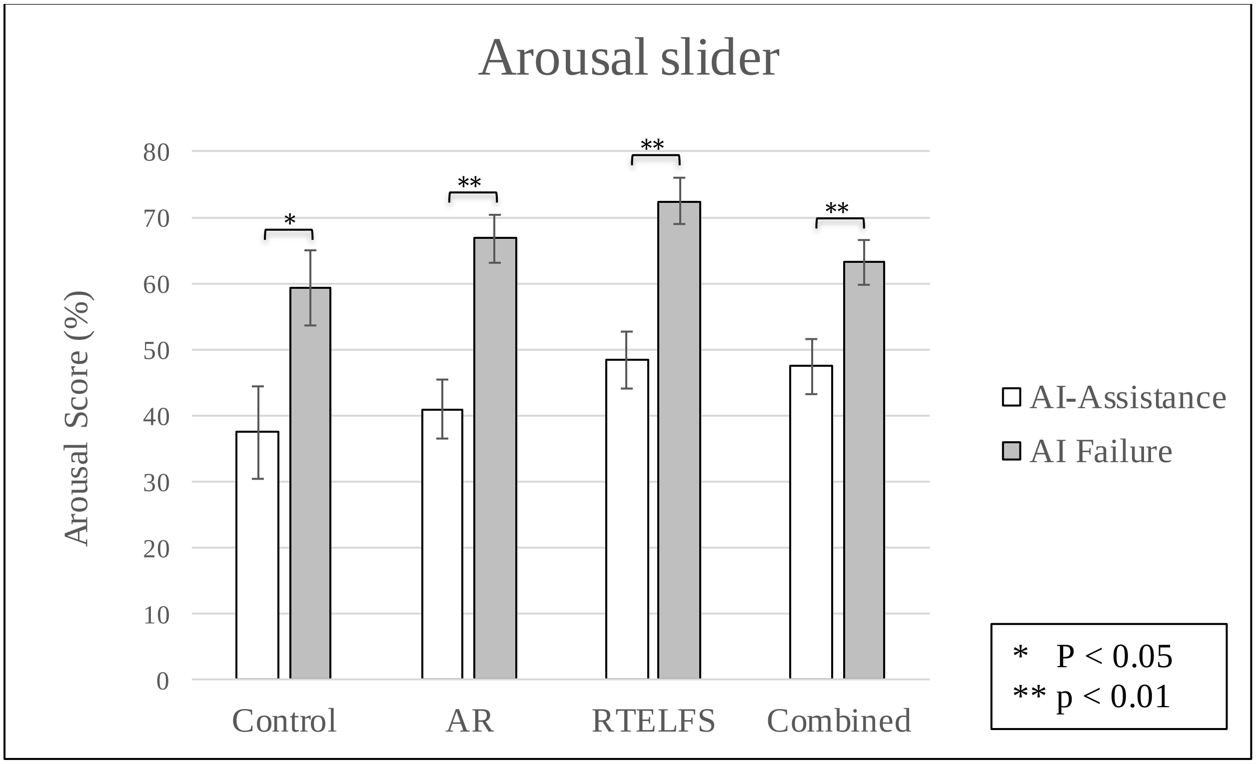

Figure 3.13 – Mean Arousal Results Across Groups for Both Tasks

3.5.4 Cognitive Engagement

Cognitive Engagement During the AI-assisted task

In the AI-assisted task, absorption scores were 3.52 for the control group, 3.52 in the Ar group, 3,22 in the Bs group, and 3.57 in the combined group (see **Figure 3.14**). None of the comparisons between the countermeasures groups and the control group showed statistically significant differences in absorption scores during the AI-assisted task. Detailed statistical values for each comparison are provided in **Table 3.1**.

Resilience of Cognitive Engagement After AI Failure (Manual Task)

After AI had failed, all groups reported increases in absorption. However, these differences were not significant. The control group reported a 0.64 increase, the AR group a 0.11 increase, the BS group a 0.47 increase and the combined group a 0.09 increase (see **Figure 3.14**). However, these increases were not significantly different in the countermeasure groups compared to the control group.

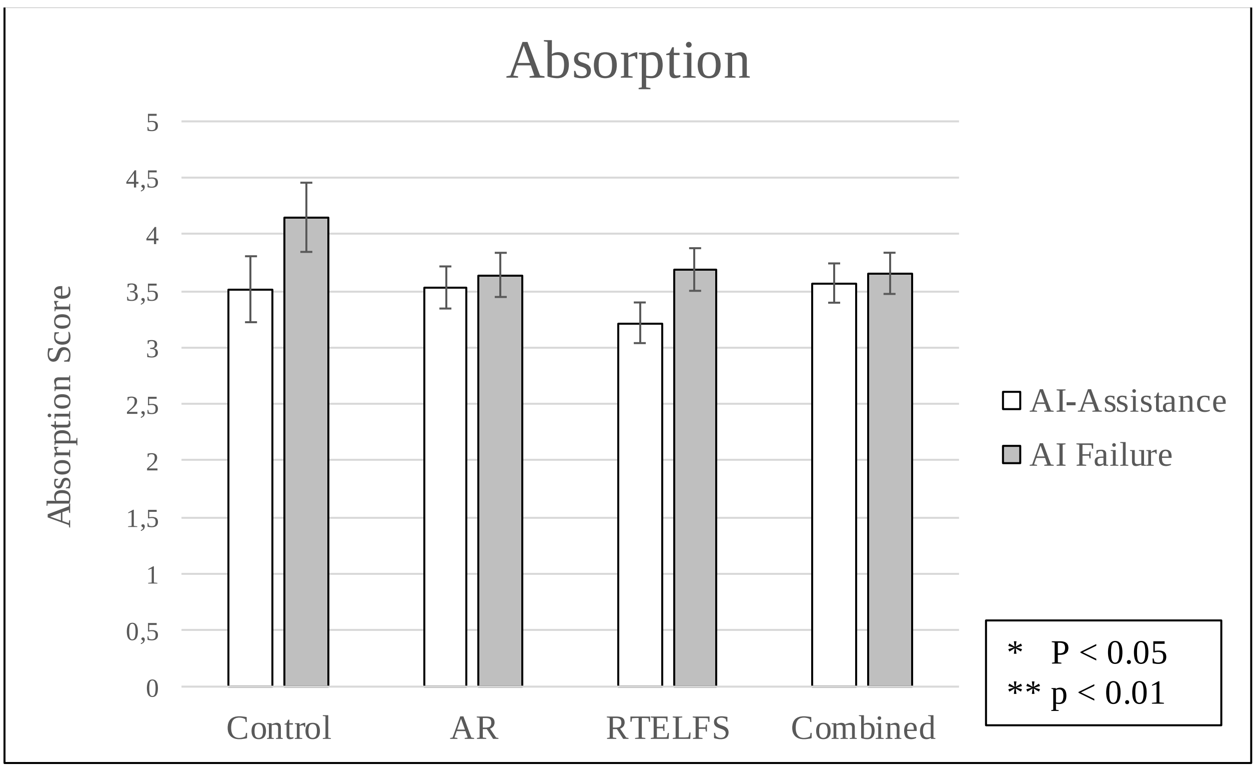

Figure 3.14 – Mean UWES Absorption Subscale Results Across Groups for Both Tasks

3.5.5 Motivation

Motivation During the AI-assisted task

During the AI-assisted task, the intrinsic motivation results were 2.75 in the control group, 3.22 in the AR group, 3.20 in the BS group, and 2.95 in the combined countermeasure group (see **Figure 3.15**). Although the intrinsic motivation was not significantly different in the countermeasures groups compared to the control group, results show that intrinsic motivation was in average higher in the countermeasure groups than in the control group. Compared to the control group, intrinsic motivation was, on average, 9.4% higher in the AR group, 9% higher in the BS group, and 4% in the combined group.

Resilience of Motivation After AI Failure (Manual Task)

After the AI failed, the results of intrinsic motivation were 3.27 for the control, 3.13 for AR, 3.22 for BS, and 2.81 for combined (see **Figure 3.15**). This shows mixed effects on intrinsic motivation after the AI failed. Increases in intrinsic motivation were reported in the control group and in the BS group, respectively, of 0.52 and 0.02, while decreases were reported in the AR and combined group, respectively, of 0.09 and 0.15. However, these effects were not significantly different in the countermeasure groups compared to the control group, as detailed in Table 3.3.

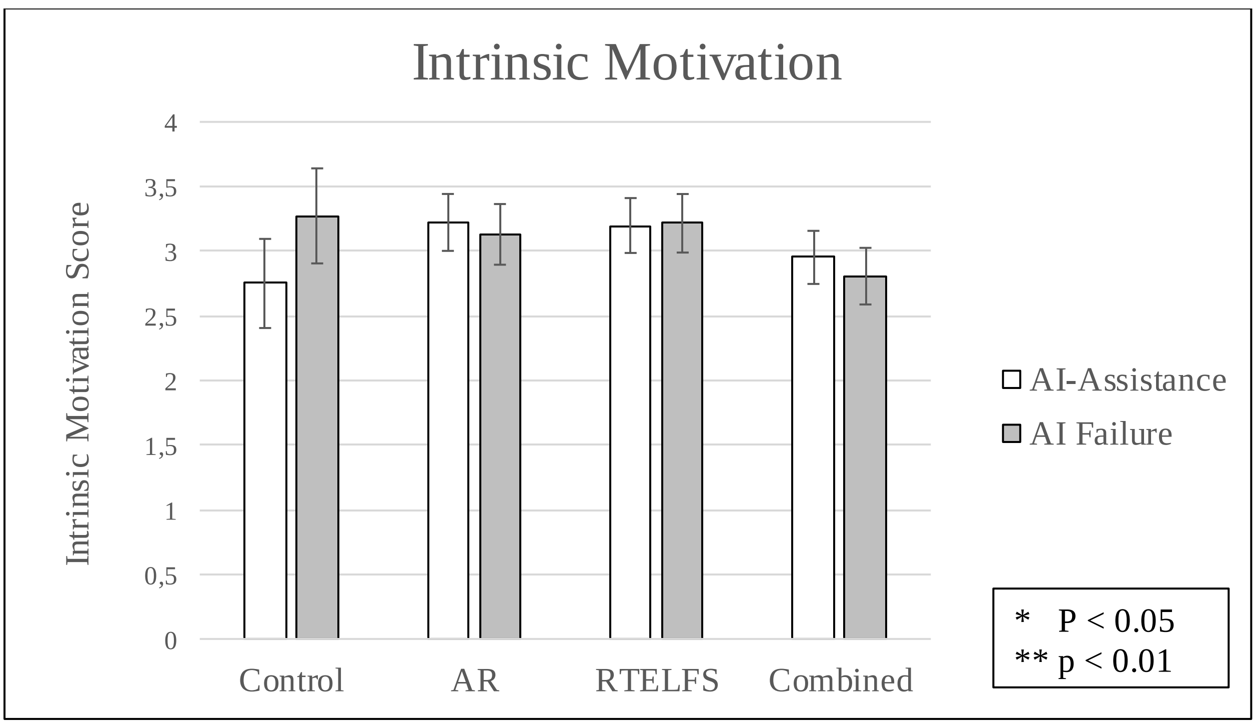

Figure 3.15 – Mean Intrinsic Motivation Results Across Groups for Both Tasks

3.6 Discussion

The objective of this study was to evaluate the effect of two cognitive countermeasures (i.e., AR and RTELFS) on the performance, engagement, and motivation of operators during AI-assisted tasks, as well as on their resilience when they had to take over the automated process. Our main finding is that both countermeasures led to significantly increased resilience of performance when the AI assistance was removed, thus supporting H2e. This was supported by manipulation performance data, which showed a decrease in the operator's manipulation performance of 8.5% between task 1 and task 2 in the Control group, compared to decreases of only 1.9%, 1.4%, and 2.3% in the Countermeasures groups (respectively the AR, RTELFS, and Combined groups). One reason that could explain this increased manipulation performance resilience is that participants may have been more engaged during the AI-assisted task when using the countermeasures, which would have led to better information retention and performance when they had to take over the automated process. This would be consistent with Karran et al. (2019), who found positive impacts of RTELFS on operator performance and levels of sustained attention, as well as with Yang et al. (2023), who showed increased learning when using AR systems. No other resilience effects were observed for motivation, cognitive engagement, emotional engagement, or behavioral engagement, failing to confirm H2a, H2b, H2c, and H2d.

The explanation that countermeasures led to higher engagement in the AI-assisted task aligns with the results from the RTELFS group. Participants in the RTELFS group reported significantly higher acceleration during the AI-assisted task, indicating that this countermeasure helped them be more physically and behaviorally engaged in the task, thus supporting H1c. Moreover, they completed the AI-assisted task in average 103 seconds faster than the control group, which, although not significant, represents a 12.5% increase in productivity. One possible explanation for this could be attributed to a potential bias in participants' understanding of how the RTELFS worked. Without a clear understanding of what "engaged" meant, participants might have thought that being more active would improve results, thus increasing their behavioral engagement and possibly turning it into a competition or game that helped increase their productivity. Additionally, although the difference was not significant, participants in the RTELFS group generally reported higher emotional valence and arousal compared to Control, indicating a potential trend that

RTELFS could enhance emotional engagement in AI-assisted tasks. This may be confirmed with larger datasets. However, in this study, these results do not support H1c.

In contrast, for AR, no statistical differences were found in engagement metrics during the AI-assisted task, failing to support H1a, H1b, and H1c. This creates a certain ambiguity in explaining why skill resilience improved for participants that used AR. One potential reason for this lack of significant effect on engagement is that simply using AR to make information more accessible and dynamic is not enough to promote worker engagement. To enhance worker engagement, especially in passive-monitoring contexts, other strategies, such as gamification, could be employed (Nguyen and Meixner, 2019). Despite this result, the increased resilience of manual performance may be attributed to the increased accessibility of information that could have improved operator learning during the AI-assisted task. This would be consistent with the findings of Yang et al. (2023) who showed increased learning outcomes and information retention when training with AR systems. Additionally, although the difference was not significant, we found that participants reported generally higher emotional valence and arousal compared to the control group. This could indicate a small increase in emotional engagement, suggesting that participants using AR were enjoying the task more than those in the control group. This would be consistent with the findings of Nguyen and Meixner (2019) who showed increased engagement in a gamified AR-assisted manufacturing context.

Surprisingly, we found no effect of the countermeasures on cognitive engagement during the AI-assisted task. For AR, this result contrasts with the findings of Nguyen and Meixner (2019) and may be attributed to the fact that the system employed in this study was a projection system rather than a head-mounted display, creating less of an immersive experience. For RTELFS, this finding aligns with Karran et al. (2019), who found no significant differences in subjective cognitive load assessments with RTELFS despite improvements in objective measures of sustained attention. One possible explanation, supported by Karran et al. (2019), is that participants were potentially unaware of the benefits of modulating their engagement levels with the RTELFS countermeasure, resulting in no subjective effect. Combining subjective and objective measurements could have provided a more comprehensive understanding of the impact on cognitive engagement. However, present findings only indicate that both countermeasures had no effect on the cognitive engagement of operators, failing to confirm H1a.

Despite slightly higher levels of intrinsic motivation in all the countermeasure groups, no significant differences in motivation were found in countermeasure groups during the AI-assisted task, failing to confirm H1d. This may be attributed to the fact that participants were students and not actual manufacturing operators, which could have affected the extent to which they relate to the work they were doing, a crucial component for motivation.

One interesting finding is that during the AI-assisted task, manipulation performance was very high across all groups, with no significant differences between them (99.09% for the Control group, 99.14% for the AR group, 98.97% for the BF group, and 98.82% for the combined group). This high level of performance was expected, as the AI system guided participants with 100% reliability. However, the lack of differences between conditions indicates that the integration of countermeasures did not distract or negatively impact the manual performance of operators. This suggests that additional technological support, such as countermeasures, could be implemented during AI-assisted work without affecting human operators' manual performance.

No specific hypothesis was established regarding the combined effect of the countermeasures. However, one potential prediction could have been that their combination would enhance outcomes, as suggested by the theory of multimedia learning (Mayer and Moreno, 1998), which posits that combining different media can improve learning outcomes. Contrary to this expectation, our findings revealed that the combination of countermeasures did not lead to improved effects. In fact, we observed slightly lower emotional valence and motivation compared to the use of single countermeasures alone. One reason for this may be attributed to the fact that both countermeasures relied on visual signals, which could have led to interference between the two technologies.

Table 3.4 Summary of Supported Hypotheses

Hypothesis	Supported ?
H1: The use of cognitive countermeasures will enhance engagement, motivation, and performance during AI-assisted work.	
H1a: During AI-assisted work, the use of cognitive will lead to greater cognitive engagement compared to individuals who did not use countermeasures.	Not supported
H1b: During AI-assisted work, the use of cognitive countermeasures will lead to greater emotional engagement compared to individuals who did not use countermeasures.	Not supported
H1c: During AI-assisted work, the use of cognitive countermeasures will lead to	Supported for

greater behavioral engagement compared to individuals who did not use countermeasures.	RTELFS
H1d: During AI-assisted work, the use of cognitive countermeasures will lead to greater motivation compared to individuals who did not use countermeasures.	Not supported
H1e: During AI-assisted work, the use of cognitive countermeasures will lead to greater performance compared to individuals who did not use countermeasures.	Not supported
H2: When automation is removed, workers that used the countermeasures would demonstrate greater resilience than those in the control group.	
H2a: When automation is removed, workers that used the countermeasures will maintain their cognitive engagement better than those who did not	Not supported
H2b: When automation is removed, workers that used the countermeasures will maintain their emotional engagement better than those who did not.	Not supported
H2c: When automation is removed, workers that used the countermeasures will maintain their behavioral engagement better than those who did not.	Not supported
H2d: When automation is removed, workers that used the countermeasures will maintain their motivation better than those who did not.	Not supported
H2e: When automation is removed, workers that used the countermeasures will maintain their performance better than those who did not.	Supported

3.7 Conclusion

In this study, we evaluated the impact of two cognitive countermeasures (i.e., AR and RTELFS) on manufacturing operators' engagement, motivation, performance, and resilience during AI-assisted work. Our findings indicate that AR and RTELFS could be used to increase operators' skill resilience in situations where operators must regain manual control of AI-automated processes. This has implications for technology designers, as it shows the potential of complementing automation systems with cognitive countermeasures to help mitigate potential human performance issues caused by automation. Moreover, we saw that RTELFS specifically enhanced the behavioral engagement of operators during AI-assisted work and led to a general increase in productivity. This demonstrates that RTELFS could be utilized to keep manufacturing operators behaviorally engaged, which would be particularly useful in training and potentially certain safety-critical settings. Surprisingly, no significant effects of the countermeasures were found on cognitive engagement, emotional engagement, and motivation, which shows that although these countermeasures helped increase performance, they had limited impact on the psychosocial human factors. Future research should, therefore, try to understand how technological aids could help increase human psychosocial factors, such as motivation and engagement, to help increase human-machine interaction in smart manufacturing contexts.

It is worth to mention some limitations to this study. First, most participants recruited in this study were engineering students and not real manufacturing operators, which could impact the extent to which these results would be applicable in real assembly contexts. Second, because the experimental procedure was conducted in a single session with only a short interval between tasks, it limits our capacity to project these results for long-term performance improvements. Third, cognitive engagement was only measured subjectively, which significantly reduces the precision and accuracy of this measure. Future work should aim to assess this dimension using both objective and subjective methods.

References

Atici-Ulusu, H., Ikiz, Y. D., Taskapilioglu, O., & Gunduz, T. (2021). Effects of augmented reality glasses on the cognitive load of assembly operators in the automotive industry. International Journal of Computer Integrated Manufacturing, 34(5), 487-499. https://doi.org/10.1080/0951192X.2021.1901314

Bakker, A. B., & Demerouti, E. (2008). Towards a model of work engagement. Career development international, 13(3), 209-223.

Biondi, F. N., Saberi, B., Graf, F., Cort, J., Pillai, P., & Balasingam, B. (2023). Distracted worker: Using pupil size and blink rate to detect cognitive load during manufacturing tasks. Applied Ergonomics, 106, 103867. https://doi.org/https://doi.org/10.1016/j.apergo.2022.103867

Büschel, W., Mitschick, A., & Dachselt, R. (2018). Here and now: Reality-based information retrieval: Perspective paper. Proceedings of the 2018 Conference on Human Information Interaction & Retrieval,

Casner, S. M., & Schooler, J. W. (2015). Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. Consciousness and Cognition, 35, 33-41.

Cooke, F. L., Cooper, B., Bartram, T., Wang, J., & Mei, H. (2019). Mapping the relationships between high-performance work systems, employee resilience and engagement: a study of the banking industry in China. The International Journal of Human Resource Management, 30(8), 1239-1260. https://doi.org/10.1080/09585192.2015.1137618

Couture, L., Passalacqua, M., Joblot, L., Magnani, F., Pellerin, R., & Léger, P.-M. (2024). Adaptive System to Enhance Operator Engagement during Smart Manufacturing Work. Sensors & Transducers., 265(2), 106-119.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies: why and how. Proceedings of the 1st international conference on Intelligent user interfaces,

Dehais, F., Causse, M., & Tremblay, S. (2011). Mitigation of conflicts with automation: use of cognitive countermeasures. Human Factors, 53(5), 448-460.

Dehais, F., Lafont, A., Roy, R., & Fairclough, S. (2020). A neuroergonomics approach to mental workload, engagement and human performance. Frontiers in Neuroscience, 14, 268. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7154497/pdf/fnins-14-00268.pdf

Dehais, F., Tessier, C., Christophe, L., & Reuzeau, F. (2010, 2010//). The Perseveration Syndrome in the Pilot's Activity: Guidelines and Cognitive Countermeasures. Human Error, Safety and Systems Development, Berlin, Heidelberg.

Demazure, T., Karran, A., Léger, P.-M., Labonté-LeMoyne, É., Sénécal, S., Fredette, M., & Babin, G. (2021). Enhancing Sustained Attention. Business & Information Systems Engineering, 63(6), 653-668. https://doi.org/10.1007/s12599-021-00701-3

Endsley, M. R. (2018). Level of automation forms a key aspect of autonomy design. Journal of Cognitive Engineering and Decision Making, 12(1), 29-34.

Endsley, M. R. (2023). Ironies of artificial intelligence. Ergonomics, 66(11), 1656-1668. https://doi.org/10.1080/00140139.2023.2243404

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.

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. Journal of Personality and Social Psychology, 69(1), 153-166. https://doi.org/10.1037/0022-3514.69.1.153

Fredricks, J., Blumenfeld, P., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. Review of Educational Research, 74, 109 - 159.

Gao, X., & Feng, H. (2023). AI-Driven Productivity Gains: Artificial Intelligence and Firm Productivity. Sustainability, 15(11), 8934. https://www.mdpi.com/2071-1050/15/11/8934

Hallberg, U. E., & Schaufeli, W. B. (2006). "Same same" but different? Can work engagement be discriminated from job involvement and organizational commitment? European psychologist, 11(2), 119-127.

He, F., Yuan, L., Mu, H., Ros, M., Ding, D., Pan, Z., & Li, H. (2023). Research and application of artificial intelligence techniques for wire arc additive manufacturing: a state-of-the-art review. Robotics and Computer-Integrated Manufacturing, 82, 102525. https://doi.org/https://doi.org/10.1016/j.rcim.2023.102525

Herrmann, F. (2018). The smart factory and its risks. Systems, 6(4), 38.

Karran, A. J., Demazure, T., Leger, P.-M., Labonte-LeMoyne, E., Senecal, S., Fredette, M., & Babin, G. (2019). Toward a Hybrid Passive BCI for the Modulation of Sustained Attention Using EEG and fNIRS [Original Research]. Frontiers in Human Neuroscience, 13. https://doi.org/10.3389/fnhum.2019.00393

Lam, M. C., Sadik, M. J., & Elias, N. F. (2021). The effect of paper-based manual and stereoscopic-based mobile augmented reality systems on knowledge retention. Virtual Reality, 25(1), 217-232. https://doi.org/10.1007/s10055-020-00451-9

Latta, G. F., & Fait, J. I. (2016). Sources of motivation and work engagement: a cross-industry analysis of differentiated profiles. Journal of Organizational Psychology, 16(2).

Liu, C., Chen, B., Guo, Y., Chen, X., Cruz Nacpil, E. J., Hou, W., & Zheng, R. (2024). Evaluation of a user-accessible countermeasure: Effect of manual dexterity gymnastics on passive driver fatigue. Transportation Research Part F: Traffic Psychology and Behaviour, 101, 387-402. https://doi.org/https://doi.org/10.1016/j.trf.2024.01.012

Madni, A. M., & Jackson, S. (2009). Towards a conceptual framework for resilience engineering. IEEE Systems Journal, 3(2), 181-191.

Mangler, J., Diwol, K., Etz, D., Rinderle-Ma, S., Ferscha, A., Reiner, G., Kastner, W., Bougain, S., Pollak, C., & Haslgrübler, M. (2021). Sustainability Through Cognition Aware Safety Systems--Next Level Human-Machine-Interaction. arXiv preprint arXiv:2110.07003.

Manikandan, N., Thejasree, P., Vimal, K., Sivakumar, K., & Kiruthika, J. (2023). Applications of Artificial Intelligence Tools in Advanced Manufacturing. In Industry 4.0 Technologies: Sustainable Manufacturing Supply Chains: Volume II-Methods for transition and trends (pp. 29-42). Springer.

Mariza, I. (2016). The impact of employees' motivation and engagement on employees' performance of manufacturing companies in Jakarta Indonesia. International Journal of Applied Business and Economic Research, 14(15), 10611-10628.

Mattera, G., Nele, L., & Paolella, D. (2024). Monitoring and control the Wire Arc Additive Manufacturing process using artificial intelligence techniques: a review. Journal of Intelligent Manufacturing, 35(2), 467-497.

Mayer, R. E., & Moreno, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. Journal of educational psychology, 91(2), 358-368.

Mazzetti, G., Robledo, E., Vignoli, M., Topa, G., Guglielmi, D., & Schaufeli, W. B. (2021). Work Engagement: A meta-Analysis Using the Job Demands-Resources Model. Psychological Reports, 126(3), 1069-1107. https://doi.org/10.1177/00332941211051988

Moray, N., & Inagaki, T. (2000). Attention and complacency. Theoretical Issues in Ergonomics Science, 1(4), 354-365.

Mypati, O., Mukherjee, A., Mishra, D., Pal, S. K., Chakrabarti, P. P., & Pal, A. (2023). A critical review on applications of artificial intelligence in manufacturing. Artificial Intelligence Review, 56(1), 661-768. https://doi.org/10.1007/s10462-023-10535-y

Naderpour, M., Nazir, S., & Lu, J. (2015). The role of situation awareness in accidents of large-scale technological systems. Process Safety and Environmental Protection, 97, 13-24. https://doi.org/https://doi.org/10.1016/j.psep.2015.06.002

Nguyen, D., & Meixner, G. (2019, 1-4 Sept. 2019). Gamified Augmented Reality Training for An Assembly Task: A Study About User Engagement. 2019 Federated Conference on Computer Science and Information Systems (FedCSIS),

Nti, I. K., Adekoya, A. F., Weyori, B. A., & Nyarko-Boateng, O. (2022). Applications of artificial intelligence in engineering and manufacturing: a systematic review. Journal of Intelligent Manufacturing, 33(6), 1581-1601.

Ojo, A. O., Fawehinmi, O., & Yusliza, M. Y. (2021). Examining the Predictors of Resilience and Work Engagement during the COVID-19 Pandemic. Sustainability, 13(5), 2902. https://www.mdpi.com/2071-1050/13/5/2902

Ong, S. K., Yuan, M. L., & Nee, A. Y. C. (2008). Augmented reality applications in manufacturing: a survey. International Journal of Production Research, 46(10), 2707-2742. https://doi.org/10.1080/00207540601064773

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.

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. Psychological Bulletin, 116(2), 220.

Passalacqua, M., Pellerin, R., Yahia, E., Magnani, F., Rosin, F., Joblot, L., & Léger, P.-M. (2024). Practice with less AI makes perfect: partially automated AI during training leads to better worker motivation, engagement, and skill acquisition. International Journal of Human–Computer Interaction, 1-21.

Peruzzini, M., & Pellicciari, M. (2017). A framework to design a human-centred adaptive manufacturing system for aging workers. Advanced Engineering Informatics, 33, 330-349. https://doi.org/https://doi.org/10.1016/j.aei.2017.02.003

Plathottam, S. J., Rzonca, A., Lakhnori, R., & Iloeje, C. O. (2023). A review of artificial intelligence applications in manufacturing operations. Journal of Advanced Manufacturing and Processing, 5(3), e10159.

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. Biological Psychology, 40(1-2), 187-195. https://www.sciencedirect.com/science/article/pii/0301051195051163?via%3Dihub

Pourabdollahian, B., Taisch, M., & Kerga, E. (2012). Serious games in manufacturing education: Evaluation of learners' engagement. Procedia Computer Science, 15, 256-265.

Raj, M., & Seamans, R. (2018). Artificial intelligence, labor, productivity, and the need for firm-level data. In The economics of artificial intelligence: An agenda (pp. 553-565). University of Chicago Press.

Romero, D., & Stahre, J. (2021). Towards the resilient operator 5.0: The future of work in smart resilient manufacturing systems. Procedia cirp, 104, 1089-1094.

Romero, D., Stahre, J., Wuest, T., Noran, O., Bernus, P., Fast-Berglund, Å., & Gorecky, D. (2016). Towards an operator 4.0 typology: a human-centric perspective on the fourth industrial revolution technologies. proceedings of the international conference on computers and industrial engineering (CIE46), Tianjin, China,

Runji, J. M., Lee, Y.-J., & Chu, C.-H. (2023). Systematic literature review on augmented reality-based maintenance applications in manufacturing centered on operator needs. International Journal of Precision Engineering and Manufacturing-Green Technology, 10(2), 567-585.

Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford publications.

Sahu, C. K., Young, C., & Rai, R. (2021). Artificial intelligence (AI) in augmented reality (AR)-assisted manufacturing applications: a review. International Journal of Production Research, 59(16), 4903-4959. https://doi.org/10.1080/00207543.2020.1859636

Saint-Lot, J., Imbert, J.-P., & Dehais, F. (2020). Red alert: a cognitive countermeasure to mitigate attentional tunneling. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,

Saks, A. M. (2006). Antecedents and consequences of employee engagement. Journal of managerial psychology, 21(7), 600-619.

Schaufeli, W. B. (2013). What is engagement? In Employee engagement in theory and practice (pp. 15-35). Routledge.

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms [Review]. Frontiers in public health, 5. https://doi.org/10.3389/fpubh.2017.00258

Sheridan, T. B. (2012). Human Supervisory Control. In Handbook of human factors and ergonomics (pp. 990-1015). https://doi.org/https://doi.org/10.1002/9781118131350.ch34

Skinner, B. F. (2019). The behavior of organisms: An experimental analysis. BF Skinner Foundation.

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing Cognitive Distraction in the Automobile. Human Factors, 57(8), 1300-1324. https://doi.org/10.1177/0018720815575149

Tortorella, G. L., Powell, D., Hines, P., Mac Cawley Vergara, A., Tlapa-Mendoza, D., & Vassolo, R. (2024). How does artificial intelligence impact employees' engagement in lean organisations? International Journal of Production Research, 1-17.

Van den Broeck, A., Howard, J. L., Van Vaerenbergh, Y., Leroy, H., & Gagné, M. (2021). Beyond intrinsic and extrinsic motivation: A meta-analysis on self-determination theory's multidimensional conceptualization of work motivation. Organizational Psychology Review, 11(3), 240-273.

Wang, Z., Bai, X., Zhang, S., Billinghurst, M., He, W., Wang, P., Lan, W., Min, H., & Chen, Y. (2022). A comprehensive review of augmented reality-based instruction in manual assembly, training and repair. Robotics and Computer-Integrated Manufacturing, 78, 102407.

Werrlich, S., Eichstetter, E., Nitsche, K., & Notni, G. (2017). An overview of evaluations using augmented reality for assembly training tasks. International Journal of Computer and Information Engineering, 11(10), 1068-1074.

Yamamoto, I. (2019). The impact of AI and information technologies on worker stress. VoxEU, https://voxeu. org/article/impact-ai-and-information-technologies-worker-stress.[40].

Yang, T., Yi, X., Lu, S., Johansson, K. H., & Chai, T. (2021). Intelligent Manufacturing for the Process Industry Driven by Industrial Artificial Intelligence. Engineering, 7(9), 1224-1230. https://doi.org/https://doi.org/10.1016/j.eng.2021.04.023

Yang, X., Mao, W., Hu, Y., Wang, J., Wan, X., & Fang, H. (2023). Does augmented reality help in industrial training? A comprehensive evaluation based on natural human behavior and knowledge retention. International Journal of Industrial Ergonomics, 98, 103516. https://doi.org/https://doi.org/10.1016/j.ergon.2023.103516

Yung, M., Kolus, A., Wells, R., & Neumann, W. P. (2020). Examining the fatigue-quality relationship in manufacturing. Applied Ergonomics, 82, 102919. https://doi.org/https://doi.org/10.1016/j.apergo.2019.102919

Zecca, G., Györkös, C., Becker, J., Massoudi, K., de Bruin, G. P., & Rossier, J. (2015). Validation of the French Utrecht Work Engagement Scale and its relationship with personality traits and impulsivity. European review of applied psychology, 65(1), 19-28.

Chapitre 4

Conclusion

4.1 Retour sur les objectifs

L'objectif de ce travail était double. Premièrement nous souhaitions proposer un nouveau système de rétroaction du niveau d'engagement spécialement conçu pour le milieu manufacturier, ce qui était l'objet du chapitre 2. Ensuite, nous voulions évaluer l'effet de deux contre-mesures cognitives (c'est-à-dire la RA et le SRNE) sur l'engagement, la motivation et la performance des opérateurs lors de tâches assistées par l'IA, ainsi que sur leur résilience dans des situations où ils devaient faire la gestion d'exceptions (telle que lorsque l'automatisation fait défaut). Ceci était l'objet du chapitre 3.

En utilisant les différences physiologiques mesurées lors d'un scénario manufacturier à haut engagement et un scénario à faible engagement, nous avons pu créer une métrique d'engagement basée sur les données de respiration et d'accélération pour évaluer le niveau d'engagement des opérateurs. Cela a permis de développer un SRNE capable de mesurer en temps réel le niveau d'engagement des opérateurs, et de classifier les niveaux d'engagement dans nos données d'entraînement avec une précision de 80,95%. Bien que le système développé puisse encore bénéficier d'améliorations en ce qui concerne la précision de sa métrique d'engagement, les résultats obtenus étaient entièrement suffisants pour une utilisation dans une étude en laboratoire. Nous pouvons donc affirmer que notre premier objectif a été atteint.

Nous avons ensuite créé une simulation manufacturière inspirée d'une visite dans une véritable usine de fabrication de raquettes de neige. Dans cette simulation, nous avons introduit une tâche de contrôle qualité et d'assemblage de 30 raquettes, qui devait être réalisée deux fois par nos participants : une première fois avec le support d'un système d'IA et une seconde fois sans aucun support, simulant une situation où le système d'IA faisait défaut. Pour cette étude, 114 participants ont été recrutés et répartis en quatre groupes, déterminant quelles contre-mesures cognitives les participants allaient recevoir lors de la tâche automatisée (RA, SRNE, la combinaison des deux, ou aucune contre-mesure). Nos résultats indiquent que les contre-mesures ont amélioré la résilience de la performance des opérateurs lorsque le support automatisé a été retiré. Plus spécifiquement, le SRNE a mené à un plus grand engagement comportemental dans la tâche assistée, comparativement au groupe contrôle. Nous pouvons donc également affirmer que notre second objectif a été atteint.

4.2 Principaux résultats

Le principal résultat à retenir de ce mémoire est que l'utilisation de contre-mesures cognitives (c'est-à-dire la RA et le SRNE) a amélioré la performance des opérateurs lorsque nous avons retiré l'assistance de l'IA. Ceci suggère que le fait d'employer la RA ou le SRNE dans un environnement de travail manufacturier assisté par l'IA permettrait aux opérateurs de développer des compétences plus résilientes, améliorant ainsi leur performance lors d'échecs d'automatisation. À notre surprise, les contre-mesures n'ont pas eu d'effet sur l'engagement cognitif ni sur la motivation des opérateurs. Cependant, elles ont entraîné un engagement émotionnel légèrement supérieur à celui du groupe de contrôle, bien que non significatif. Additionnellement, lors de la tâche assistée par l'IA, le SRNE, a mené à un plus grand engagement physique et comportemental et à une amélioration de la productivité de 12,5%.

En ce qui concerne le développement du SRNE dans le chapitre 2, nous avons été capables de concevoir un nouvel index d'engagement basé sur les différences physiologiques entre un scénario manufacturier hautement engageant et un scénario faiblement engageant. En utilisant cet indice dans une étape de validation, nous avons été capables de prédire les niveaux d'engagement de notre base de données d'entraînement avec une précision de 80.25%.

4.3 Contributions

Ce travail a trois principales contributions. Premièrement, il propose un nouveau système de rétroaction du niveau d'engagement qui est adapté au milieu manufacturier. L'avantage de ce système est qu'il utilise des outils de mesure physiologiques comme la respiration et l'accélération qui sont faciles à collecter en contexte manufacturier où les opérateurs sont souvent en mouvement. Contrairement aux systèmes similaires présentés dans la littérature, qui reposent principalement sur l'électroencéphalographie, ce système est plus pratique pour un environnement manufacturier. Cependant, ce système a été développé pour un contexte spécifique, et les futurs développements devront s'assurer de l'adapter avant de l'implémenter.

La méthodologie que nous avons employée pour développer le SRNE peut également représenter une contribution. En effet, malgré le fait que ce n'est pas la première fois que des gens simulent des scénarios plus ou moins engageants pour construire des métriques (voir Moray et Inagaki, 2000; Verdière et al., 2018), c'est la première fois qu'un système de rétroaction a été conçu en utilisant cette méthodologie. Ainsi, des chercheurs pourraient utiliser la méthodologie proposée dans le chapitre 2 pour développer des systèmes qui sont adaptés à leur contexte, que ce soit en aviation, en conduite autonome ou en manufacturier.

Finalement, nos résultats indiquent que l'utilisation de contre-mesures cognitives peut aider à mitiger certains effets négatifs de l'implantation d'IA en manufacturier, notamment sur la performance des opérateurs manufacturiers. Ceci peut aider les développeurs de technologies industrielles à améliorer l'interaction humain-machine de leurs produits. De plus, cela peut fournir aux dirigeants d'usines intelligentes de nouvelles pistes de solutions pour assurer la productivité de leurs opérations. À notre connaissance, ceci est le premier travail qui explore les effets de contre-mesures cognitives comme moyen de mitiger les potentiels effets négatifs de l'IA sur les opérateurs manufacturiers.

4.4 Futures Recherches

Les futures recherches devraient essayer de comprendre par quels mécanismes la RA et le SRNE ont permis d'aider les opérateurs à conserver leurs compétences. Pour la RA, il se pourrait que ce soit dû à la plus grande accessibilité de l'information ou par l'interaction de l'information sur l'environnement réel, potentiellement induisant des comportements différents chez les opérateurs. De plus, puisque l'utilisation de ludification a démontré un grand potentiel dans la littérature pour rehausser l'engagement (Nguyen and Meixner, 2019), il serait intéressant d'évaluer si la ludification à travers une application de réalité augmentée pourrait rehausser l'engagement perçu des opérateurs. Le SRNE devrait être amélioré pour tenir compte des différences physiologiques entre les individus et comparer sa fiabilité à des mesures d'engagement plus établies dans la littérature comme l'EEG et la dilatation de la pupille. De plus, il serait intéressant d'évaluer si l'effet positif des contre-mesures permet d'améliorer la rétention 'habileté au long terme.

Bibliographie

Allan Cheyne, J., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. Cognition, 111(1), 98-113. https://doi.org/https://doi.org/10.1016/j.cognition.2008.12.009

Argyle, E. M., Marinescu, A., Wilson, M. L., Lawson, G., & Sharples, S. (2021). Physiological indicators of task demand, fatigue, and cognition in future digital manufacturing environments. International Journal of Human-Computer Studies, 145, 102522. https://doi.org/https://doi.org/10.1016/j.ijhcs.2020.102522

Atchley, P., Dressel, J., Jones, T. C., Burson, R. A., & Marshall, D. (2011). Talking and driving: applications of crossmodal action reveal a special role for spatial language. Psychological research, 75, 525-534. https://link.springer.com/content/pdf/10.1007/s00426-011-0342-7.pdf

Atici-Ulusu, H., Ikiz, Y. D., Taskapilioglu, O., & Gunduz, T. (2021). Effects of augmented reality glasses on the cognitive load of assembly operators in the automotive industry. International Journal of Computer Integrated Manufacturing, 34(5), 487-499. https://doi.org/10.1080/0951192X.2021.1901314

Bakker, A. B., & Demerouti, E. (2008). Towards a model of work engagement. Career development international, 13(3), 209-223.

Benarroch, E. E. (2009). The locus ceruleus norepinephrine system: functional organization and potential clinical significance. Neurology, 73(20), 1699-1704.

Bernabei, M., & Costantino, F. (2024). Adaptive automation: Status of research and future challenges. Robotics and Computer-Integrated Manufacturing, 88, 102724. https://doi.org/https://doi.org/10.1016/j.rcim.2024.102724

Biondi, F. N., Saberi, B., Graf, F., Cort, J., Pillai, P., & Balasingam, B. (2023). Distracted worker: Using pupil size and blink rate to detect cognitive load during manufacturing tasks. Applied Ergonomics, 106, 103867. https://doi.org/https://doi.org/10.1016/j.apergo.2022.103867

Büschel, W., Mitschick, A., & Dachselt, R. (2018). Here and now: Reality-based information retrieval: Perspective paper. Proceedings of the 2018 Conference on Human Information Interaction & Retrieval,

Carter, B. T., & Luke, S. G. (2020). Best practices in eye tracking research. International Journal of Psychophysiology, 155, 49-62. https://doi.org/https://doi.org/10.1016/j.ijpsycho.2020.05.010

Casner, S. M., & Schooler, J. W. (2015). Vigilance impossible: Diligence, distraction, and daydreaming all lead to failures in a practical monitoring task. Consciousness and Cognition, 35, 33-41.

Castiblanco Jimenez, I. A., Gomez Acevedo, J. S., Marcolin, F., Vezzetti, E., & Moos, S. (2023). Towards an integrated framework to measure user engagement with interactive or physical products. International Journal on Interactive Design and Manufacturing (IJIDeM), 17(1), 45-67.

Cherif, N. H., Mezghani, N., Gaudreault, N., Ouakrim, Y., Mouzoune, I., & Boulay, P. (2018). Physiological data validation of the hexoskin smart textile.

Cooke, F. L., Cooper, B., Bartram, T., Wang, J., & Mei, H. (2019). Mapping the relationships between high-performance work systems, employee resilience and engagement: a study of the banking industry in China. The International Journal of Human Resource Management, 30(8), 1239-1260. https://doi.org/10.1080/09585192.2015.1137618

Couture, L., Passalacqua, M., Joblot, L., Magnani, F., Pellerin, R., & Léger, P.-M. (2024). Adaptive System to Enhance Operator Engagement during Smart Manufacturing Work. Sensors & Transducers., 265(2), 106-119.

Couture, L., Passalacqua, M., Joblot, L., Magnani, F., Pellerin, R., & Léger, P.-M. (2024). Enhancing Operator Engagement during AI-assisted Manufacturing Work Using Optimal State Deviation Feedback System.

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies: why and how. Proceedings of the 1st international conference on Intelligent user interfaces, de Guinea, A. O., Titah, R., & Léger, P.-M. (2013). Measure for measure: A two study multi-trait multi-method investigation of construct validity in IS research. Computers in Human Behavior, 29(3), 833-844.

de Guinea, A. O., Titah, R., & Léger, P.-M. (2014). Explicit and implicit antecedents of users' behavioral beliefs in information systems: A neuropsychological investigation. Journal of Management Information Systems, 30(4), 179-210.

Dehais, F., Causse, M., & Tremblay, S. (2011). Mitigation of conflicts with automation: use of cognitive countermeasures. Human Factors, 53(5), 448-460.

Dehais, F., Dupres, A., Di Flumeri, G., Verdiere, K., Borghini, G., Babiloni, F., & Roy, R. (2018). Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI. 2018 IEEE international conference on systems, man, and cybernetics (SMC),

Dehais, F., Lafont, A., Roy, R., & Fairclough, S. (2020). A Neuroergonomics Approach to Mental Workload, Engagement and Human Performance [Review]. Frontiers in Neuroscience, 14. https://doi.org/10.3389/fnins.2020.00268

Dehais, F., Tessier, C., Christophe, L., & Reuzeau, F. (2010, 2010//). The Perseveration Syndrome in the Pilot’s Activity: Guidelines and Cognitive Countermeasures. Human Error, Safety and Systems Development, Berlin, Heidelberg.

Demazure, T., Karran, A., Léger, P.-M., Labonté-LeMoyne, É., Sénécal, S., Fredette, M., & Babin, G. (2021). Enhancing Sustained Attention. Business & Information Systems Engineering, 63(6), 653-668. https://doi.org/10.1007/s12599-021-00701-3

Eldenfria, A., & Al-Samarraie, H. (2019). Towards an online continuous adaptation mechanism (OCAM) for enhanced engagement: An EEG study. International Journal of Human–Computer Interaction, 35(20), 1960-1974.

Endsley, M. R. (2018). Level of automation forms a key aspect of autonomy design. Journal of Cognitive Engineering and Decision Making, 12(1), 29-34.

Endsley, M. R. (2023). Ironies of artificial intelligence. Ergonomics, 66(11), 1656-1668. https://doi.org/10.1080/00140139.2023.2243404

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.

Fairclough, S. H., & Venables, L. (2005). Psychophysiological predictors of task engagement and distress. Human Factors in Design, Safety, and Management In D. de Waard, KA Brookhuis, R. van Egmond, and Th. Boersema (Eds.), 349-362.

Feigh, K. M., Dorneich, M. C., & Hayes, C. C. (2012). Toward a Characterization of Adaptive Systems:A Framework for Researchers and System Designers. Human Factors, 54(6), 1008-1024. https://doi.org/10.1177/0018720812443983

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. Journal of Personality and Social Psychology, 69(1), 153-166. https://doi.org/10.1037/0022-3514.69.1.153

Fredricks, J., Blumenfeld, P., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. Review of Educational Research, 74, 109 - 159.

Gao, X., & Feng, H. (2023). AI-Driven Productivity Gains: Artificial Intelligence and Firm Productivity. Sustainability, 15(11), 8934. https://www.mdpi.com/2071-1050/15/11/8934

Goujon, A., Rosin, F., Magnani, F., Lamouri, S., Pellerin, R., & Joblot, L. (2024). Industry 5.0 use cases development framework. International Journal of Production Research, 1-26.

Gouraud, J., Delorme, A., & Berberian, B. (2018). Out of the Loop, in Your Bubble: Mind Wandering Is Independent From Automation Reliability, but Influences Task Engagement [Original Research]. Frontiers in Human Neuroscience, 12. https://doi.org/10.3389/fnhum.2018.00383

Hajra, S. G., Xi, P., & Law, A. (2020, 11-14 Oct. 2020). A comparison of ECG and EEG metrics for in-flight monitoring of helicopter pilot workload. 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Hallberg, U. E., & Schaufeli, W. B. (2006). “Same same” but different? Can work engagement be discriminated from job involvement and organizational commitment? European psychologist, 11(2), 119-127.

Hansen, A. L., Johnsen, B. H., & Thayer, J. F. (2003). Vagal influence on working memory and attention. International Journal of Psychophysiology, 48(3), 263-274.

He, D., Wang, Z., Khalil, E. B., Donmez, B., Qiao, G., & Kumar, S. (2022). Classification of Driver Cognitive Load: Exploring the Benefits of Fusing Eye-Tracking and Physiological Measures. Transportation Research Record, 2676(10), 670-681. https://doi.org/10.1177/03611981221090937

He, F., Yuan, L., Mu, H., Ros, M., Ding, D., Pan, Z., & Li, H. (2023). Research and application of artificial intelligence techniques for wire arc additive manufacturing: a state-of-the-art review. Robotics and Computer-Integrated Manufacturing, 82, 102525. https://doi.org/https://doi.org/10.1016/j.rcim.2023.102525

Herrmann, F. (2018). The smart factory and its risks. Systems, 6(4), 38.

Hinss, M. F., Brock, A. M., & Roy, R. N. (2022). Cognitive effects of prolonged continuous human-machine interaction: The case for mental state-based adaptive interfaces [Review]. Frontiers in Neuroergonomics, 3. https://doi.org/10.3389/fnrgo.2022.935092

Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. J. (2015). The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. Biological Psychology, 110, 100-106. https://doi.org/https://doi.org/10.1016/j.biopsycho.2015.06.013

Karran, A. J., Demazure, T., Leger, P.-M., Labonte-LeMoyne, E., Senecal, S., Fredette, M., & Babin, G. (2019). Toward a Hybrid Passive BCI for the Modulation of Sustained Attention Using EEG and fNIRS [Original Research]. Frontiers in Human Neuroscience, 13. https://doi.org/10.3389/fnhum.2019.00393

Körber, M., Cingel, A., Zimmermann, M., & Bengler, K. (2015). Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving. Procedia Manufacturing, 3, 2403-2409. https://doi.org/https://doi.org/10.1016/j.promfg.2015.07.499

Kundinger, T., Sofra, N., & Riener, A. (2020). Assessment of the Potential of Wrist-Worn Wearable Sensors for Driver Drowsiness Detection. Sensors, 20(4), 1029. https://www.mdpi.com/1424-8220/20/4/1029

https://mdpi-res.com/d_attachment/sensors/sensors-20-01029/article_deploy/sensors-20-01029-v2.pdf?version=1582299430

Lam, M. C., Sadik, M. J., & Elias, N. F. (2021). The effect of paper-based manual and stereoscopic-based mobile augmented reality systems on knowledge retention. Virtual Reality, 25(1), 217-232. https://doi.org/10.1007/s10055-020-00451-9

Latta, G. F., & Fait, J. I. (2016). Sources of motivation and work engagement: a cross-industry analysis of differentiated profiles. Journal of Organizational Psychology, 16(2).

Léger, P.-M., Courtemanche, F., Fredette, M., & Sénécal, S. (2019). A cloud-based lab management and analytics software for triangulated human-centered research. Information Systems and Neuroscience: NeuroIS Retreat 2018,

Léger, P.-M., Davis, F. D., Cronan, T. P., & Perret, J. (2014). Neurophysiological correlates of cognitive absorption in an enactive training context. Computers in Human Behavior, 34, 273-283. https://doi.org/https://doi.org/10.1016/j.chb.2014.02.011

Léger, P.-M., Karran, A. J., Courtemanche, F., Fredette, M., Tazi, S., Dupuis, M., Hamza, Z., Fernández-Shaw, J., Côté, M., & Del Aguila, L. (2022). Caption and observation based on the algorithm for triangulation (COBALT): Preliminary results from a beta trial. In NeuroIS Retreat (pp. 229-235). Springer.

Li, B.-h., Hou, B.-c., Yu, W.-t., Lu, X.-b., & Yang, C.-w. (2017). Applications of artificial intelligence in intelligent manufacturing: a review. Frontiers of Information Technology & Electronic Engineering, 18(1), 86-96.

Liu, C., Chen, B., Guo, Y., Chen, X., Cruz Nacpil, E. J., Hou, W., & Zheng, R. (2024). Evaluation of a user-accessible countermeasure: Effect of manual dexterity gymnastics on passive driver fatigue. Transportation Research Part F: Traffic Psychology and Behaviour, 101, 387-402. https://doi.org/https://doi.org/10.1016/j.trf.2024.01.012

Lu, Y., Zheng, H., Chand, S., Xia, W., Liu, Z., Xu, X., Wang, L., Qin, Z., & Bao, J. (2022). Outlook on human-centric manufacturing towards Industry 5.0. Journal of Manufacturing Systems, 62, 612-627.

Madni, A. M., & Jackson, S. (2009). Towards a conceptual framework for resilience engineering. IEEE Systems Journal, 3(2), 181-191.

Mangler, J., Diwol, K., Etz, D., Rinderle-Ma, S., Ferscha, A., Reiner, G., Kastner, W., Bougain, S., Pollak, C., & Haslgrübler, M. (2021). Sustainability Through Cognition Aware Safety Systems--Next Level Human-Machine-Interaction. arXiv preprint arXiv:2110.07003.

Manikandan, N., Thejasree, P., Vimal, K., Sivakumar, K., & Kiruthika, J. (2023). Applications of Artificial Intelligence Tools in Advanced Manufacturing. In Industry 4.0 Technologies: Sustainable Manufacturing Supply Chains: Volume II-Methods for transition and trends (pp. 29-42). Springer.

Mariza, I. (2016). The impact of employees' motivation and engagement on employees' performance of manufacturing companies in Jakarta Indonesia. International Journal of Applied Business and Economic Research, 14(15), 10611-10628.

Mattera, G., Nele, L., & Paolella, D. (2024). Monitoring and control the Wire Arc Additive Manufacturing process using artificial intelligence techniques: a review. Journal of Intelligent Manufacturing, 35(2), 467-497.

Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., Grier, R., & Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry. Emotion, 2(4), 315.

Mayer, R. E., & Moreno, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. Journal of educational psychology, 91(2), 358-368.

Mazzetti, G., Robledo, E., Vignoli, M., Topa, G., Guglielmi, D., & Schaufeli, W. B. (2021). Work Engagement: A meta-Analysis Using the Job Demands-Resources Model. Psychological Reports, 126(3), 1069-1107. https://doi.org/10.1177/00332941211051988

Mccraty, R., & Shaffer, F. (2015). Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-regulatory Capacity, and Health Risk. Global Advances in Health and Medicine, 4(1), 46-61. https://doi.org/10.7453/gahmj.2014.073

Moray, N., & Inagaki, T. (2000). Attention and complacency. Theoretical Issues in Ergonomics Science, 1(4), 354-365.

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell R, G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. Psychophysiology, 48(11), 1532-1543. https://doi.org/10.1111/j.1469-8986.2011.01226.x

Mypati, O., Mukherjee, A., Mishra, D., Pal, S. K., Chakrabarti, P. P., & Pal, A. (2023). A critical review on applications of artificial intelligence in manufacturing. Artificial Intelligence Review, 56(1), 661-768. https://doi.org/10.1007/s10462-023-10535-y

Naderpour, M., Nazir, S., & Lu, J. (2015). The role of situation awareness in accidents of large-scale technological systems. Process Safety and Environmental Protection, 97, 13-24. https://doi.org/https://doi.org/10.1016/j.psep.2015.06.002

Naujoks, F., Höfling, S., Purucker, C., & Zeeb, K. (2018). From partial and high automation to manual driving: Relationship between non-driving related tasks, drowsiness and take-over performance. Accident Analysis & Prevention, 121, 28-42. https://doi.org/https://doi.org/10.1016/j.aap.2018.08.018

Nguyen, D., & Meixner, G. (2019, 1-4 Sept. 2019). Gamified Augmented Reality Training for An Assembly Task: A Study About User Engagement. 2019 Federated Conference on Computer Science and Information Systems (FedCSIS),

Nti, I. K., Adekoya, A. F., Weyori, B. A., & Nyarko-Boateng, O. (2022). Applications of artificial intelligence in engineering and manufacturing: a systematic review. Journal of Intelligent Manufacturing, 33(6), 1581-1601.

Ojo, A. O., Fawehinmi, O., & Yusliza, M. Y. (2021). Examining the Predictors of Resilience and Work Engagement during the COVID-19 Pandemic. Sustainability, 13(5), 2902. https://www.mdpi.com/2071-1050/13/5/2902

Ong, S. K., Yuan, M. L., & Nee, A. Y. C. (2008). Augmented reality applications in manufacturing: a survey. International Journal of Production Research, 46(10), 2707-2742. https://doi.org/10.1080/00207540601064773

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. (2000). Designing automation for human use: empirical studies and quantitative models. Ergonomics, 43(7), 931-951. https://doi.org/10.1080/001401300409125

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance Consequences of Automation-Induced 'Complacency'. The International Journal of Aviation Psychology, 3(1), 1-23. https://doi.org/10.1207/s15327108ijap0301_1

Pashler, H. (1994). Dual-task interference in simple tasks: data and theory. Psychological Bulletin, 116(2), 220.

Passalacqua, M., Cabour, G., Pellerin, R., Léger, P.-M., & Doyon-Poulin, P. (2024). Human-centered AI for industry 5.0 (HUMAI5. 0): Design framework and case studies. In Human-centered AI (pp. 260-274). Chapman and Hall/CRC.

Passalacqua, M., Léger, P.-M., Nacke, L. E., Fredette, M., Labonté-Lemoyne, É., Lin, X., Caprioli, T., & Sénécal, S. (2020). Playing in the backstore: interface gamification increases warehousing workforce engagement. Industrial Management & Data Systems, 120(7), 1309-1330.

Passalacqua, M., Pellerin, R., Yahia, E., Magnani, F., Rosin, F., Joblot, L., & Léger, P.-M. (2024). Practice with less AI makes perfect: partially automated AI during training leads to better worker motivation, engagement, and skill acquisition. International Journal of Human–Computer Interaction, 1-21.

Peruzzini, M., & Pellicciari, M. (2017). A framework to design a human-centred adaptive manufacturing system for aging workers. Advanced Engineering Informatics, 33, 330-349. https://doi.org/https://doi.org/10.1016/j.aei.2017.02.003

Plathottam, S. J., Rzonca, A., Lakhnori, R., & Iloeje, C. O. (2023). A review of artificial intelligence applications in manufacturing operations. Journal of Advanced Manufacturing and Processing, 5(3), e10159.

Pooladvand, S., & Hasanzadeh, S. (2023). Impacts of Stress on Workers' Risk-Taking Behaviors: Cognitive Tunneling and Impaired Selective Attention. Journal of Construction Engineering and Management, 149(8), 04023060. https://doi.org/doi:10.1061/JCEMD4.COENG-13339

Pope, A. T., Bogart, E. H., & Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. Biological Psychology, 40(1-2), 187-195. https://www.sciencedirect.com/science/article/pii/0301051195051163?via%3Dihub

Pourabdollahian, B., Taisch, M., & Kerga, E. (2012). Serious games in manufacturing education: Evaluation of learners' engagement. Procedia Computer Science, 15, 256-265.

Raj, M., & Seamans, R. (2018). Artificial intelligence, labor, productivity, and the need for firm-level data. In The economics of artificial intelligence: An agenda (pp. 553-565). University of Chicago Press.

Riedl, R., Fischer, T., Léger, P.-M., & Davis, F. D. (2020). A decade of NeuroIS research: progress, challenges, and future directions. ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 51(3), 13-54.

Romero, D., & Stahre, J. (2021). Towards the resilient operator 5.0: The future of work in smart resilient manufacturing systems. Procedia cirp, 104, 1089-1094.

Romero, D., Stahre, J., Wuest, T., Noran, O., Bernus, P., Fast-Berglund, Å., & Gorecky, D. (2016). Towards an operator 4.0 typology: a human-centric perspective on the fourth industrial revolution technologies. proceedings of the international conference on computers and industrial engineering (CIE46), Tianjin, China,

Rosin, F., Forget, P., Lamouri, S., & Pellerin, R. (2021). Impact of Industry 4.0 on decision-making in an operational context. Advances in Production Engineering & Management, 16(4).

Rosin, F., Forget, P., Lamouri, S., & Pellerin, R. (2022). Enhancing the decision-making process through industry 4.0 technologies. Sustainability, 14(1), 461.

Roy, R. N., Bovo, A., Gateau, T., Dehais, F., & Carvalho Chanel, C. P. (2016). Operator Engagement During Prolonged Simulated UAV Operation. IFAC-PapersOnLine, 49(32), 171-176. https://doi.org/https://doi.org/10.1016/j.ifacol.2016.12.209

Runji, J. M., Lee, Y.-J., & Chu, C.-H. (2023). Systematic literature review on augmented reality-based maintenance applications in manufacturing centered on operator needs. International Journal of Precision Engineering and Manufacturing-Green Technology, 10(2), 567-585.

Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford publications.

Sahu, C. K., Young, C., & Rai, R. (2021). Artificial intelligence (AI) in augmented reality (AR)-assisted manufacturing applications: a review. International Journal of Production Research, 59(16), 4903-4959. https://doi.org/10.1080/00207543.2020.1859636

Saint-Lot, J., Imbert, J.-P., & Dehais, F. (2020). Red alert: a cognitive countermeasure to mitigate attentional tunneling. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems,

Saks, A. M. (2006). Antecedents and consequences of employee engagement. Journal of managerial psychology, 21(7), 600-619.

Scerbo, M. (2007). Adaptive automation. Neuroergonomics: The brain at work, 239252.

Schaufeli, W. B. (2013). What is engagement? In Employee engagement in theory and practice (pp. 15-35). Routledge.

Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2003). Utrecht work engagement scale-9. Educational and Psychological Measurement.

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms [Review]. Frontiers in public health, 5. https://doi.org/10.3389/fpubh.2017.00258

Shaffer, F., McCraty, R., & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. Frontiers in Psychology, 5, 1040. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4179748/pdf/fpsyg-05-01040.pdf

Sheridan, T. B. (2012). Human Supervisory Control. In Handbook of human factors and ergonomics (pp. 990-1015). https://doi.org/https://doi.org/10.1002/9781118131350.ch34

Skinner, B. F. (2019). The behavior of organisms: An experimental analysis. BF Skinner Foundation.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. Psychological Bulletin, 132(6), 946-958. https://doi.org/10.1037/0033-2909.132.6.946

Soni, R., & Muniyandi, M. (2019). Breath Rate Variability: A Novel Measure to Study the Meditation Effects. International Journal of Yoga, 12(1), 45-54. https://doi.org/10.4103/ijoy.IJOY_27_17

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing Cognitive Distraction in the Automobile. Human Factors, 57(8), 1300-1324. https://doi.org/10.1177/0018720815575149

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. Perspectives on Psychological Science, 10(1), 82-96. https://journals.sagepub.com/doi/10.1177/1745691614556681?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub 0pubmed

Tortorella, G. L., Powell, D., Hines, P., Mac Cawley Vergara, A., Tlapa-Mendoza, D., & Vassolo, R. (2024). How does artificial intelligence impact employees' engagement in lean organisations? International Journal of Production Research, 1-17.

Vadeboncoeur, D., Pellerin, R., & Danjou, C. (2024). Assessing the influence of human factors on overall labor effectiveness in manufacturing: a comprehensive literature review. Automation, Robotics & Communications for Industry 4.0/5.0, 135.

Van den Broeck, A., Howard, J. L., Van Vaerenbergh, Y., Leroy, H., & Gagné, M. (2021). Beyond intrinsic and extrinsic motivation: A meta-analysis on self-determination theory's multidimensional conceptualization of work motivation. Organizational Psychology Review, 11(3), 240-273.

Vasseur, A., Passalacqua, M., Sénécal, S., & Léger, P.-M. (2023). The Use of Eye-tracking in Information Systems Research: A Literature Review of the Last Decade. AIS Transactions on Human-Computer Interaction, 15(3), 292-321.

Verdière, K. J., Roy, R. N., & Dehais, F. (2018). Detecting Pilot's Engagement Using fNIRS Connectivity Features in an Automated vs. Manual Landing Scenario [Original Research]. Frontiers in Human Neuroscience, 12. https://doi.org/10.3389/fnhum.2018.00006

Wang, Z., Bai, X., Zhang, S., Billinghurst, M., He, W., Wang, P., Lan, W., Min, H., & Chen, Y. (2022). A comprehensive review of augmented reality-based instruction in manual assembly, training and repair. Robotics and Computer-Integrated Manufacturing, 78, 102407.

Werrlich, S., Eichstetter, E., Nitsche, K., & Notni, G. (2017). An overview of evaluations using augmented reality for assembly training tasks. International Journal of Computer and Information Engineering, 11(10), 1068-1074.

Wientjes, C. J. E. (1992). Respiration in psychophysiology: methods and applications. Biological Psychology, 34(2), 179-203. https://doi.org/https://doi.org/10.1016/0301-0511(92)90015-M

Williams, D. P., Thayer, J. F., & Koenig, J. (2016). Resting cardiac vagal tone predicts intraindividual reaction time variability during an attention task in a sample of young and healthy adults. Psychophysiology, 53(12), 1843-1851.

Yamamoto, I. (2019). The impact of AI and information technologies on worker stress. VoxEU, https://voxeu. org/article/impact-ai-and-information-technologies-worker-stress.[40].

Yang, T., Yi, X., Lu, S., Johansson, K. H., & Chai, T. (2021). Intelligent Manufacturing for the Process Industry Driven by Industrial Artificial Intelligence. Engineering, 7(9), 1224-1230. https://doi.org/https://doi.org/10.1016/j.eng.2021.04.023

Yang, X., Mao, W., Hu, Y., Wang, J., Wan, X., & Fang, H. (2023). Does augmented reality help in industrial training? A comprehensive evaluation based on natural human behavior and knowledge retention. International Journal of Industrial Ergonomics, 98, 103516. https://doi.org/https://doi.org/10.1016/j.ergon.2023.103516

Yung, M., Kolus, A., Wells, R., & Neumann, W. P. (2020). Examining the fatigue-quality relationship in manufacturing. Applied Ergonomics, 82, 102919. https://doi.org/https://doi.org/10.1016/j.apergo.2019.102919

Yurish, S. (2024). Proceedings of the 4th IFSA Winter Conference on Automation, Robotics and Communications for Industry 4.0/5.0 (ARCI 2024). https://doi.org/10.13140/RG.2.2.20923.18722

Zecca, G., Györkös, C., Becker, J., Massoudi, K., de Bruin, G. P., & Rossier, J. (2015). Validation of the French Utrecht Work Engagement Scale and its relationship with personality traits and impulsivity. European review of applied psychology, 65(1), 19-28.

Annexe A

Acte de Conférence présenté à la conférence ARCI2024 (7 au 9 février 2024)

Oral ☒ / Poster ☐ / The same ☐

Topic: Special session - Integrating Human Factors in Operation Management Models

Enhancing Operator Engagement During AI-Assisted Manufacturing Work Using Optimal State Deviation Feedback System

L. Couture [1], **M. Passalacqua** [2], **L. Joblot** [3], **F. Magnani** [4], **R. Pellerin** [2] **and P. M. Léger** [1]
[1] HEC Montréal, Tech3Lab, 3000 Chemin de la Côte-Sainte-Catherine, H3T 2A7, Montréal, Canada
[2] Polytechnique Montréal, 2500 Chemin de Polytechnique, H3T 1J4, Montréal, Canada
[3] Arts et Métiers Institute of Technology, LISPEN / HESAM University / UBFC, Chalon-Sur-Saône, France
[4] Université Aix-Marseille, CERGAM, Centrale Méditerranée, 38, rue Frédéric Joliot-Curie, 13013, Marseille, France
Tel.: + 15149981557, e-mail: loic.couture@hec.ca

Summary: The integration of Artificial Intelligence (AI) in manufacturing is shifting the focus of operators from manual labor to cognitive supervision roles. While this transition demands more engagement from operators, the less stimulating nature of monitoring tasks has, paradoxically, reduced operator involvement, consequently presenting new challenges in performance maintenance. Addressing this issue, our research adopted an iterative design science methodology to create a biocybernetic system that aims to enhance operator engagement in their evolving workplace. This system leverages physiological signals to intuitively display how much an operator's engagement level deviates from an ideal state, ensuring operators stay aware of their psychophysiological state of engagement and can quickly adjust to any decreases in engagement. In this paper, we detail the 4-step process that led to the development of the first version of the system. Capitalizing on the physiological differences observed in manufacturing operators during "high" and "low" engagement scenarios, we defined a task-specific Optimal State Deviation Index (OSDI) formula. This formula enabled us to predict participants' engagement states with an 80.95% success rate in our testing dataset.

Keywords: Biocybernetic system, Manufacturing, Engagement, Automation, Design Science, Artificial Intelligence

1. Introduction

AI-driven automation is transforming manufacturing operators' roles, shifting their work from manual work to supervising systems [1], which can lead to less stimulating tasks, adversely impacting operator engagement and performance [2]. Given the prevalent risk of occupational injuries associated with manufacturing work [3], it appears essential for operators to maintain an optimal engagement state. Specifically, operators must avoid excessive vigilance, which can increase fatigue over time or lead to cognitive tunneling, a state in which operators adopt a narrow focus and neglect other important information [4]. Operators must also avoid cognitive underload, which can result in mind wandering and inattention [5]. In addressing the issue of maintaining an optimal engagement state, very little research has explored how new technologies can effectively improve operator engagement in manufacturing.

However, the work of Demazure et al. [6] is particularly promising in this regard. Their research demonstrated the potential of using real-time engagement level feedback to significantly improve users' attentiveness. Our study seeks to adapt this approach for manufacturing, aiming to develop a tool to help operators maintain optimal engagement levels.

The structure of this paper is outlined as follows. Section 2 delves into the reasons for developing a new biocybernetic system tailored for manufacturing. Section 3 is dedicated to detailing the iterative design science methodology that was employed to create the system. The results that influenced the system's design are detailed in Section 4. Finally, we present our concluding remarks, along with a discussion of the current system's limitations in Section 5.

2. Background

The integration of automated systems offers significant advantages for industrial applications. Nonetheless, it is important to acknowledge that most of these automated systems have yet to achieve perfection in terms of system reliability [7]. Consequently, in instances where a system's reliability is not absolute, it is prudent for enterprises to deploy human operators. These operators play a crucial role in monitoring automated systems' functionality, enabling the early detection of anomalies and facilitating timely intervention to rectify such occurrences. However, monitoring automated systems can present several human challenges, including a decrease in vigilance over time [8] and low monitoring performance [2]. This decline is attributed to both cognitive overload, which can result in cognitive fatigue and cognitive tunneling [9], and cognitive underload, which can cause mind wandering, low motivation, and increased distraction [10]. Therefore, one way to tackle this issue is to ensure the operator can balance their level of engagement throughout the monitoring task.

In this context, the work of Demazure et al. [6] seems particularly promising as it offers a passive system that informs the operator of their level of engagement in real-time. This feature not only keeps operators aware of their mental state in real-time but also enables them to make immediate adjustments as needed. The solution proposed by Demazure utilizes electroencephalography (EEG) signals to provide a real-time, intuitive display of the operator's engagement level through a color gradient display. Karran et al. [11] employed Demazure et al.'s system and demonstrated that continuously showing engagement levels to operators notably enhanced sustained attention during long monitoring tasks. This was evidenced by increased EEG wave coherence recorded for participants who received continuous engagement feedback. In contrast, participants who did not receive engagement feedback and those who only received engagement feedback after critical disengagement thresholds were reached reported low EEG wave coherence. While these results appear encouraging, a notable challenge with this solution is the necessity of an accurate measurement of engagement, which can be particularly difficult in manufacturing settings.

Numerous physiological tools have been used in the literature to measure task engagement, including eye-tracking [12], electroencephalography (EEG) [6], electrodermal activity (EDA) [13], and heart rate variability measures (HRV) [14]. Although eye-tracking and EEG methods are well-established in the literature for assessing engagement, their practical application in manufacturing faces significant challenges. The primary issue with these techniques is their limited adaptability to the dynamic nature of manufacturing environments. Operators in such settings are frequently mobile and engage with their surroundings in a 360-degree manner. This constant movement and the need to interact with a wide-ranging environment render both eye-tracking and EEG methodologies less feasible due to their inherent requirement for relative stability and controlled observation conditions. EDA is typically measured on the palm of the hand, which could constrain operators in their work. However, HRV can accurately be assessed during operator movement, making it a potential choice for a manufacturing setting [15]. HRV is defined as the variation of time intervals between consecutive heartbeats [16] and is mainly used as a measure of the activation of the autonomous nervous system [17]. There is, however, some debate regarding the interpretation of HRV measures [17,18], which raises questions regarding the viability of using this metric to assess task engagement.

This ambiguity makes Moray and Inagaki's approach [19] particularly appealing. Their method evaluates monitoring performance by contrasting actual operator performance to an optimal standard. From this perspective, for any specific task, it seems feasible to establish a performance metric by initially recording the responses of an operator in a high-performance scenario and comparing it to a low-performance scenario. Therefore, when we want to assess operator engagement, a potential approach would be to establish an engagement metric by comparing physiological responses recorded in highly engaging scenarios with those from a minimally engaging scenario, using contrast to construct a reliable measure of engagement for this particular task. Additionally, since increasing the level of automation has been shown to be the source of lower engagement [20], it seems possible to use the levels of automation to induce different levels of engagement in a manufacturing context.

Hence, to maintain optimal engagement levels of manufacturing operators within their dynamic work environments, our proposal involves developing a new biocybernetic system inspired by the research of Demazure et al. [6] but tailored to the manufacturing context. Rather than depending on exact engagement metrics and measurements, our system follows a methodology similar to Moray and Inagaki [19], leveraging physiological indicators that differentiate between optimal and suboptimal engagement states. A significant advantage of this approach is its adaptability to complex settings like manufacturing, where constraints exist concerning the feasibility of certain physiological measurements, such as eye-tracking and EEG.

3. Methods

We used a design science methodology to develop an optimal state deviation feedback system involving a four-step process that included three studies (see **Table 1**). The first two steps were dedicated to identifying physiological markers that could characterize the reduction of operators' engagement during a specific task and developing a biocybernetic system. The last two steps were dedicated to evaluating different features of the biocybernetic system, i.e., the display modality and the scaling method.

3.1. Step 1 – Collect Data

In the first step, we collected physiological and perceptual data from participants in more and less engaging manufacturing situations. We recruited 22 students (age=21.62±3.17; men=14) for a within-subject experiment, in which they twice performed a quality control and assembly task on a simulated assembly line[2]. All participants provided a signed consent in-line with the University ethics committee (project # 2023-5058) and were compensated with the sum of 40 euros. The task, explained in more detail in [2], required participants to detect errors on partially

[2] For an overview of the experimental setup: https://youtu.be/xtcpxqcyz8k.

assembled snowshoes and complete the assembly by fixing the binding to the base at its pivot point (see **Fig.1**). In the "less engaging" condition, we automated the participants' decision-making, equipping them with a fully reliable error detection system that indicated to the operator whether or not a snowshoe had a defect. In the "more engaging" condition, participants had to manually detect errors before assembling the snowshoes. During each task, a total of 30 snowshoes had to be assembled by the participants, with six being defective. Participants realized the task once with automated support and once without automated support, with condition order being randomly assigned and counterbalanced. During the task, we collected physiological data using a Hexoskin vest [21], recording heart rate, respiratory rate, and acceleration data. We also collected perceived cognitive absorption, vigor, and dedication using the Utrecht Work Engagement Scale (UWES) [22], which was collected post-task. The raw physiological data from the Hexoskin was pre-processed and synchronized using the COBALT Photobooth software [23]. The list of physiological and self-reported data collected can be found in **Table 2**.

Type of data	Measure	Description
Physiological data	Beats per minute	Number of beats per minute
	SDNN	Standard deviation of NN intervals
	LF	Power of the Low-frequency band (0.04-0.15 Hz) (ms^2)
	HF	Power of the High-frequency band (0.15-0.4 Hz) (ms^2)
	LF/HF	Ratio of Low-to-High frequency power
	Breathing Rate	Number of respirations per minute
	Minute Ventilation	Respiratory volume per minute (L/min)
	Cadence	Number of steps per minute
	Motion	Norm of the 3D acceleration vector (G)
Self-reported measures	Absorption score	Perceived absorption
	Vigor score	Perceived vigor
	Dedication score	Perceived dedication

Table 1. Methodology employed to design the biocybernetic system

Step	Step 1	Step 2	Step 3	Step 4
Title	*Collect data*	*Identify markers*	*Display validation*	*Scaling validation*
Description	Study 1: Collection of Physiological Data in Scenarios with Varied Engagement Levels	Identify physiological markers of engagement and design the system	*Study 2:* Validating multiple display modalities of engagement	*Study 3:* Validating multiple index scaling methods
Experimental design	Within-subject	-	Within-subject	Between subject
Conditions	No automation Automation	-	Discrete color gradient (3 shades of color) Continuous color gradient (100 shades between green and red)	Min/Max since the beginning of the task Min/Max of training data Min=25th and Max=75th percentiles since the beginning of the task
Experimental manipulation	Manufacturing QandA and assembly tasks using snowshoes.	Feature extraction using a logistic regression model Validation with LOOCV	Fully automated manufacturing QandA and assembly tasks using images of snowshoes.	Fully automated manufacturing QandA and assembly tasks using images of snowshoes
Data	Collected physiological data (bpm, breath rate, motion) and perceived work engagement (UWES)	Task 1 and Task 2 data from step 1	10 minutes semi-directed interviews	Five questions questionnaire
Participants	22 participants	-	3 participants	10 participants

Table 2. List of collected variables

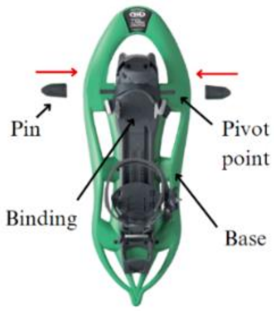

Fig. 1. Product used in the manufacturing task

3.2. Step 2 – Identify Markers and System Design

In the second step, we began by validating our primary assumption that the condition with automation was less engaging than the manual condition. Due to a noticeable learning effect between the first and second tasks, primarily manifested in performance improvements, we chose to focus exclusively on the results obtained from the first task, where no learning effects could affect perception. We compared the perceived absorption, dedication, and vigor scores between automated and manual conditions using the Mann-Whitney-Wilcoxon Test, which is suitable for comparing non-parametric independent samples.

To categorize the data, we assigned labels of "high" or "low" engagement to arrays of data, depending on the condition experienced by the participant. Data originating from the automated task was labeled as "low engagement," while data from the manual task was labeled as "high engagement". We then defined a task-specific optimal state deviation index (OSDI) using the three physiological variables with the highest estimated weights in the logistic regression model used to predict the probability of a participant being more engaged in the task. The whole dataset (Task 1 and Task 2) was used to develop the formula. The formula represents a weighted sum, where each coefficient corresponds to the respective variable's estimated power to predict if a participant is in a "high" or "low" state of engagement. The formula is based on 30-second data windows.

$$OSDI = (435.7\ \text{motion}_{\text{std}}) - (175.4\ \text{motion}_{\text{mean}}) + (0.78\ \text{breathingRate}_{\text{std}}) \quad (1)$$

To validate the formula, we employed the leave-one-out cross-validation (LOOCV) using the OSDI in a logistic model to predict if a participant's engagement during a task was "higher" or "lower". The same dataset was used for this validation step. We then developed a biocybernetic system on Python that employs the OSDI formula to calculate the index in real-time, scale it, and visually represent it as a color gradient (see **Fig.2**). The system received pre-processed physiological data every second (1 Hz) from the Hexoskin vest. It calculated the engagement index using the OSDI formula based on the last 30 seconds' data. The first prototype (and the one used for the next step) scaled the OSDI between [0-100] using the minimum and the maximum values since the beginning of the task.

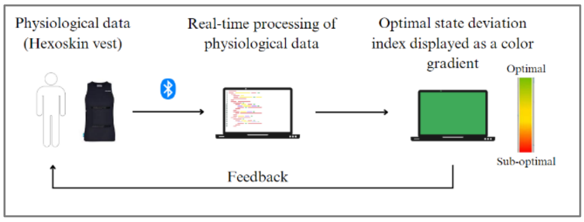

Fig 2. Overview of the biocybernetic system

3.3. Step 3 – Display Validation

In the third step, we assessed whether representing the index through a continuous color gradient (100 shades) or a discrete color gradient (3 colors) was more effective in conveying participants' engagement levels. We recruited three participants for a within-subjects pilot test. Each participant completed a low-fidelity version of the automated assembly task twice (using printed images of snowshoes instead of real snowshoes), experiencing the feedback system in both formats. After completing each task, participants underwent a 5-minute semi-directed interview. During this interview, they were asked about their perceptions of the system's impact on their engagement, the potential distractions caused by the system, and its effectiveness in representing their engagement levels. Positive and negative statements in each category were compiled and analyzed, making the decision to retain the continuous color gradient.

3.4. Step 4 – Scaling Validation

In the fourth step, we aimed to identify the most effective method for scaling the index. We tested three scaling methods: (i) dynamically adjusting the minimum and maximum values based on the minimum and maximum values recorded since the beginning of the task, (ii) using the minimum and maximum values of the training dataset, measured with formula 2 to exclude outliers, and (iii) dynamically setting the minimum and maximum values respectively to the 25th and 75th percentile of the data since the beginning of the task.

$$MIN/MAX = OSDI_{mean} \pm 3 * OSDI_{std} \quad (2)$$

We performed a between-subjects experiment with 10 participants who each completed the same low-fidelity version of the manufacturing task while being assisted by the system in one of its three possible formats (using printed images of snowshoes instead of real snowshoes). After completing the task, participants were asked to rate the representativeness, interpretability, and distractive nature of the color display on a scale from 0 to 100.

4. Results

The one-sided Mann-Whitney-Wilcoxon Test used **for step two** revealed a statistically significant difference in perceived absorption scores between manual and automated conditions ($p = .03$; $d = 0.83$), suggesting that the reported absorption scores tend to be lower in the automated condition compared to manual condition. This result supports our primary assumption that the automated condition was less engaging than the manual condition. No significant differences were found between conditions for dedication ($p = .40$; $d = -.82$) and vigor ($p = .82$; $d = -.43$) subscales of UWES during task 1 (see **Fig. 3**).

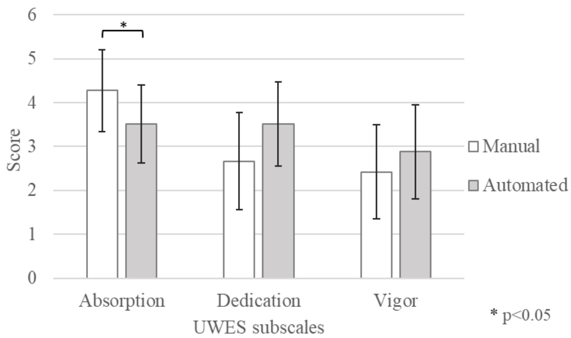

Fig. 3. Task 1 UWES Questionnaire Results: Participant Response Analysis

Using the OSDI formula to predict if a participant was in a "high" or "low" state of engagement in a logistic regression model, we achieved 81.31% accuracy on the training set and 80.95% on the testing set, as confirmed through leave-one-out cross-validation. **For step three**, where we assessed the display modality, we employed a qualitative labeling technique to categorize interview statements into three themes: effect on perceived engagement, distraction, and representativeness. The number of statements in each category was then compiled (see **Table 3**), showing that the discrete color gradient was more distracting (0 positive, six negative statements) than the continuous color gradient (2 positive, 0 negative statements).

Table 3. Compilation of qualitative statements on continuous and discrete color gradients

	Perceived effect on engagement		Distraction		Representa-tiveness	
	(+)	(-)	(+)	(-)	(+)	(-)
Continuous	5	0	2	0	2	2
Discrete	2	1	0	6	0	3

In step four, the self-reported data from questionnaires revealed that all methods were equally easy to interpret and not distracting. However, the scaling method (ii) utilizing the minimum and maximum values from the training dataset proved to be more representative, with a mean score of $93.33\% \pm 6.24\%$. This was in contrast to the scaling method (i), which was based on the minimum and maximum values since the beginning of the task ($mean = 57.33 \pm 12.28\%$), and method (iii) which was based on percentiles ($mean = 45.5 \pm 14.5\%$), as illustrated in **Fig. 4**. Based on these analyses, we concluded that the continuous color gradient and scaling method, which utilized the minimum and maximum values of the training dataset, i.e., method (ii), are preferred options for any future work.

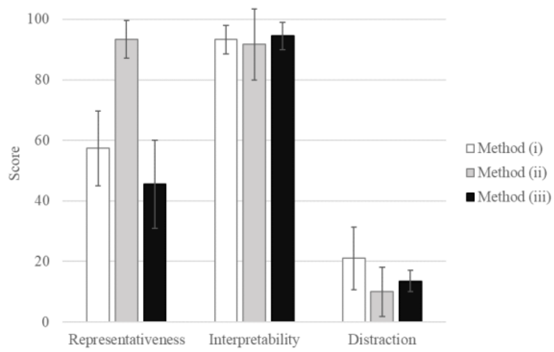

Fig. 4. Scaling method comparison: Evaluating Representativeness, Interpretability, and Distraction through Questionnaire Scores

5. Conclusions

This study employed a design science methodology to create an optimal state deviation feedback system designed to help manufacturing operators stay engaged in their workplace. The task-specific optimal state index was developed using physiological data collected during a simulated manufacturing assembly task, achieving 80.95% accuracy in predicting the engagement state of the testing set. We assessed two display modalities and three scaling methods to inform our design. The final design utilized a continuous color gradient calibrated based on the lowest and highest values of the training set. A subsequent study was conducted to test this advancement in a broader scale, which will be discussed in forthcoming scientific publications.

It is essential to acknowledge certain limitations inherent in this system. First, our assessment of engagement relied solely on self-reported data. Ideally, employing real-time physiological monitoring tools, like EEG, would have enhanced the validation of the measured engagement levels but would have been much more intrusive than the Hexoskin vest we used. Additionally, it should be noted that while the leave-out samples were not employed in training the predictive models, they were utilized in creating the OSDI formula. As a result, the model's effectiveness for new participants might not be as robust as measured in this study. Finally, it is important to note that the formula used in this system strongly depends on the task and is specifically tailored to the context of our study. This means that the OSDI formula may not yield reliable results in different contexts and, therefore, should not be applied to other scenarios without appropriate modifications and validation.

References

[1] R. Parasuraman and C. D. Wickens, "Humans: Still vital after all these years of automation," in Decision Making in Aviation: Routledge, 2017, pp. 251-260.

[2]. M. Passalacqua, R. Pellerin, E. Yahia, F. Magnani, F. Rosin, L. Joblot, P.M. Léger, Practice with Less AI Makes Perfect: Partially Automated AI during Training Leads to Better Worker Motivation, Engagement, and Skill Acquisition [Provisionally Accepted], *International Journal of Human-Computer Interaction, 2024*

[3]. Occupation groups with the highest incidence rate of nonfatal occupational injuries and illnesses* per 10,000 full-time workers in the U.S. in 2020 [Graph], Bureau of Labor Statistics, 2021. [Online].

[4] S. Pooladvand and S. Hasanzadeh, "Impacts of Stress on Workers' Risk-Taking Behaviors: Cognitive Tunneling and Impaired Selective Attention," Journal of Construction Engineering and Management, vol. 149, no. 8, p. 04023060, 2023, doi: doi:10.1061/JCEMD4.COENG-13339.

[5] F. Dehais, A. Lafont, R. Roy, and S. Fairclough, "A Neuroergonomics Approach to Mental Workload, Engagement and Human Performance," (in English), Frontiers in Neuroscience, Review vol. 14, 2020-April-07 2020, doi: 10.3389/fnins.2020.00268.

[6]. T. Demazure et al., Enhancing Sustained Attention, *Business and Information Systems Engineering*, Vol. 63, Issue 6, 2021, pp. 653-668.

[7] K. Wang, J. Lu, S. Ruan, and Y. Qi, "Continuous Error Timing in Automation: The Peak-End Effect on Human-Automation Trust," International Journal of Human–Computer Interaction, pp. 1-13, doi: 10.1080/10447318.2023.2223954.

[8] M. Körber, A. Cingel, M. Zimmermann, and K. Bengler, "Vigilance Decrement and Passive Fatigue Caused by Monotony in Automated Driving," Procedia Manufacturing, vol. 3, pp. 2403-2409, 2015/01/01/ 2015, doi: https://doi.org/10.1016/j.promfg.2015.07.499.

[9] P. A. Desmond and P. A. Hancock, "Active and passive fatigue states," in Stress, workload, and fatigue: CRC Press, 2000, pp. 455-465.

[10] S. Conte, D. Harris, and J. Blundell, "Evaluating the Impact of Passive Fatigue on Pilots Using Performance and Subjective States Measures," in International Conference on Human-Computer Interaction, 2023: Springer, pp. 21-36.

[11] A. J. Karran et al., "Toward a Hybrid Passive BCI for the Modulation of Sustained Attention Using EEG and fNIRS," (in English), Frontiers in Human Neuroscience, Original Research vol. 13, 2019-November-06 2019, doi: 10.3389/fnhum.2019.00393.

[12] R. N. Roy, A. Bovo, T. Gateau, F. Dehais, and C. P. C. Chanel, "Operator engagement during prolonged simulated uav operation," IFAC-PapersOnLine, vol. 49, no. 32, pp. 171-176, 2016.

[13] U. Kale, J. Rohács, and D. Rohács, "Operators' Load Monitoring and Management," Sensors, vol. 20, no. 17, p. 4665, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/17/4665 https://mdpi-res.com/d_attachment/sensors/sensors-20-04665/article_deploy/sensors-20-04665.pdf?version=1597827000.

[14] S. Pütz, A. Mertens, L. Chuang, and V. Nitsch, "Physiological measures of operators' mental state in supervisory process control tasks: a scoping review," Ergonomics, pp. 1-30, doi: 10.1080/00140139.2023.2289858.

[15] S. Lee, H. Kim, D.-H. Kim, M. Yum, and M. Son, "Heart rate variability in male shift workers in automobile manufacturing factories in South Korea," International Archives of Occupational and Environmental Health, vol. 88, no. 7, pp. 895-902, 2015/10/01 2015, doi: 10.1007/s00420-014-1016-8.

[16] R. Mccraty and F. Shaffer, "Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-regulatory Capacity, and Health Risk," Global Advances in Health and Medicine, vol. 4, no. 1, pp. 46-61, 2015, doi: 10.7453/gahmj.2014.073.

[17] F. Shaffer, R. McCraty, and C. L. Zerr, "A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability," Frontiers in psychology, vol. 5, p. 1040, 2014. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4179748/pdf/fpsyg-05-01040.pdf.

[18] G. E. Billman, "The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance," vol. 4, ed: Frontiers Media SA, 2013, p. 26.

[19] N. Moray and T. Inagaki, "Attention and complacency," Theoretical Issues in Ergonomics Science, vol. 1, no. 4, pp. 354-365, 2000.

[20] D. Manzey, J. Reichenbach, and L. Onnasch, "Human Performance Consequences of Automated Decision Aids: The Impact of Degree of Automation and System Experience," Journal of Cognitive Engineering and Decision Making, vol. 6, no. 1, pp. 57-87, 2012/03/01 2012, doi: 10.1177/1555343411433844.

[21]. Hexoskin. (https://www.hexoskin.com)

[22]. W. B. Schaufeli, A. B. Bakker, and M. Salanova, Utrecht work engagement scale-9, Educational and Psychological Measurement, 2003

[23] P.-M. Léger et al., "Caption and observation based on the algorithm for triangulation (COBALT): Preliminary results from a beta trial," in NeuroIS Retreat: Springer, 2022, pp. 229-23.