A DISSERTATION

In order to obtain the title of

DOCTOR OF THE UNIVERSITY OF TOULOUSE

Issued by Université Toulouse 2 - Jean Jaurès

Defended by

Noé MONSAINGEON

On 17 October 2022

Design and evaluation of multimodal interfaces guiding attentional resources to the status of partially automated driving systems

PhD School: **CLESCO - Comportement, Langage, Education, Socialisation, Cognition**

Department: **Psychologie**

Research Laboratory:
CLLE - Unité Cognition, Langues, Langage, Ergonomie

Supervised by
Céline LEMERCIER and Loïc CAROUX

The jury members

Mr. Joost DE WINTER, Referee
Mrs. Emilie LOUP-ESCANDE, Referee
Mr. Jordan NAVARRO, Examiner
Mrs. Élodie LABEYE, Examiner
Mrs. Céline LEMERCIER, PhD Supervisor
Mr. Loïc CAROUX, PhD Co-supervisor

REMERCIEMENTS

Le matin de mon deuxième jour d'école, mon père me dit qu'il faut y retourner.

« Mais Papa, j'ai fait l'école déjà, j'y suis allé hier.

- Mon fils, tu en as encore pour un petit moment. »

Il ne s'était pas trompé, car 25 ans plus tard, je sors « officiellement » de l'école. Mes parents, c'est à vous que je le dois. Un parcours chaotique aux premiers abords, et pourtant un soutien inconditionnel. Je vous remercie infiniment de m'avoir inculqué ce principe de chercher du plaisir dans le travail, car il m'a permis de trouver ma voie. Maman, merci d'avoir veillé sur ma santé (mentale et physique) pendant ces années mouvementées. Papa, merci pour ton intérêt et ta curiosité.

Merci à ma fratrie, mon frère Maël qui a fait naître en moi la curiosité pour les nouvelles technologies, ma sœur Sarah qui m'a ouvert la porte à la psychologie, mon frère Calliste dont le travail et le parcours m'ont inspiré et motivé, et ma sœur Ysoline qui m'a montré l'exemple au travers de sa ténacité et résilience.

Merci à mes directeurs de thèse. Je suis chanceux d'être tombé sur cet appel à candidatures et d'avoir été choisi. Votre présence tout au long de ce processus était indéniable et n'est pas acquise par tous les doctorants. Céline, merci pour ta proximité dès notre première rencontre. Merci de ne jamais lâcher le morceau lorsque c'est juste. Merci aussi pour ta vision et tes idées originales qui font du bien à la recherche. Tu as su, malgré un emploi du temps de fou furieux, me montrer que la passion peut toujours être présente et nous faire avancer vers de nouveaux horizons. Merci Loïc, car tu as été pour moi un maître dans la science et dans l'écriture scientifique. Merci pour ta présence et tes relectures incessantes. Mon attrait pour l'écriture a été grandement stimulé par ton aide.

Merci à mes collègues de Renault. Sabine, merci de m'avoir guidé dans ce monde nouveau de l'automobile. Tu as été une voix calme et rigoureuse dans un contexte bruyant et incertain. Jean-François, merci pour tes idées et tes digressions qui m'ont permis d'explorer des champs novateurs. Xavier, merci de défendre le point de vue de l'utilisateur dans un monde d'ingénieur. Merci à l'équipe Ergo IHM pour votre soutien, et merci à l'équipe Senso pour nos repas ensemble. Merci aux étudiants qui ont travaillé avec moi sur ce sujet passionnant, Yanna et Ghao. Vous avez été un moteur important dans l'accomplissement de ce travail. J'ai été chanceux de vous avoir à mes côtés.

Aux membres du projet CMI, Virgil, Nicolas, Clarisse, Ilia, Kahina, Jean-Marc, Alexandre, nous sommes arrivés au bout de ce projet ensemble, malgré l'incertitude sur sa finalité. Je dois bien avouer que j'ai douté pendant longtemps, mais que grâce à vous, nous y sommes arrivés.

Merci à mes collègues du CLLE et aux membres du TIMH Lab, anciennement Klaxon. Merci à mes collègues doctorants et post-doc Sharon, Maxime, Valentin, Robin, Aïsha, Eugénie, et Collin de m'avoir accueilli dans vos bureaux lorsque j'étais SBF (Sans Bureau Fixe). Un grand merci aussi de m'avoir accueilli chez vous pour des nuits de tranquillité avant des présentations stressantes. Merci également à Pierre-Vincent pour ton aide lors de mes prises de tête sur l'Eyetracking, SCANeR, et le VBA.

Ma thèse était bien affiliée aux CLLE à Toulouse, mais Grenoble reste dans mon cœur pour toujours. Le bureau E121, vous aussi m'avez accueillie dans votre bureau comme un des votre. Merci.

A Grenoble, les potos Alex, Arnaud, César, Tiff, Tommy, Margaux, et Marine (la prochaine sur la liste des futurs docteurs…) vous avez été un soutien de fou. Sans vous, sans la détente qu'on a partagée ensemble presque au quotidien, ma tête aurait explosé.

Merci mes OG Charles-Max et Geo, avec qui j'ai partagé un logement pendant ces trois années à Paname. On n'a pas eu beaucoup l'occasion de célébrer, thèse oblige, mais vous m'avez toujours soutenu et ça, ça vaut tout. Chalyzere, merci de m'avoir montré la voie de la thèse et la passion pour la recherche. Un plaisir de partager ça avec toi pendant cette étape de ma vie.

Et le meilleur pour la fin, Cynthia baby, on aura ciré ensemble les bancs de la fac du premier jour jusqu'au dernier. Quel bonheur de partager une passion et un métier comme le nôtre avec la personne que l'on aime. Cette thèse n'aurait clairement pas été possible sans toi. Tu as su m'aider, me montrer de nouvelles façons de voir les choses, me soutenir et, le plus dur, écouter mes interminables présentations. Ne parlons même pas des jours de rédaction intense où tu as tout fait pour m'aider (comme cuisiner de bons petits plats). Tout le stress que j'ai pu ressentir pendant ces 3 ans était oublié lorsque je m'endormais à tes côtés.

Et bien-sûr, à toutes les personnes qui m'ont soutenue et aidée pendant ce processus et que je n'ai pas pu citer ici, un énorme merci. Sans vous, rien de tout cela n'aurait été possible.

Writing is like driving at night in the fog. You can only see as far as your headlights, but you can make the whole trip that way.

E.L. Doctorow

RÉSUMÉ

Les systèmes d'assistance de véhicules partiellement automatisés sont capables de remplacer temporairement l'humain dans le contrôle latéral et longitudinal du véhicule. Seulement, ils peuvent rendre le contrôle du véhicule au conducteur lorsque leurs limites de fonctionnement sont atteintes. Il est donc nécessaire que les conducteurs basculent leur attention depuis la surveillance de l'activité des systèmes automatisés vers le contrôle du véhicule. Ils doivent identifier l'état de fonctionnement de ces systèmes à tout moment pour éviter une confusion. Les interfaces multimodales et de fiabilité des systèmes automatisés semblent apporter des solutions appropriées pour répondre à ces problématiques. La littérature actuelle n'est pas suffisante pour statuer sur l'efficacité de ces interfaces à induire une conscience efficace de l'état des systèmes automatisés. Cette thèse propose de concevoir et d'évaluer l'efficacité d'interfaces exploitant les modalités sonores, haptiques, et en vision périphérique à orienter l'attention des conducteurs sur l'état de fonctionnement des systèmes d'assistance. Une première étude expérimentale réalisée sur route a montré qu'une interface multimodale comportant des informations orientées sur les besoins des conducteurs permettait une meilleure compréhension du fonctionnement du véhicule qu'une interface uniquement visuelle comportant des informations orientées sur l'état du véhicule. Les deux interfaces ont provoqué des confusions de modes, et aucune n'indiquait la fiabilité des systèmes automatisés.

Deux études expérimentales ont ensuite été menées afin de concevoir et d'évaluer un indicateur d'approche des limites des systèmes automatisés perceptible en vision périphérique. Cet indicateur a été considéré comme une information utile et a permis d'anticiper des situations de suspension des systèmes automatisés. Un second groupe d'études expérimentales a été mené avec pour objectif de vérifier la capacité d'interfaces sonores et haptiques intégrée au volant à être perçues et comprises par les conducteurs. Celles-ci se sont montrées efficaces pour indiquer l'état des systèmes automatisés. Les différents éléments d'interfaces testés au préalable ont été assemblés dans un simulateur afin d'évaluer leur effet conjugué. Lors d'une étude longitudinale, des conducteurs ont été confrontés, à plusieurs reprises et aléatoirement, à quatre situations de reprise en main après suspension des systèmes automatisés (virages trop serrés, bouchons, brouillard, marquages routiers effacés). Ils possédaient soit une interface multimodale indiquant les limites des systèmes automatisés, soit une interface visuelle classique. La conscience de l'état des systèmes automatisés, la distribution de l'attention, et la confiance dans ces systèmes ont été évaluées. Les résultats montrent un effet positif de l'interface multimodale d'approche des limites sur (1) la compréhension du fonctionnement des systèmes automatisés, (2) les performances de conduite lors de la suspension des systèmes automatisés, (3) l'orientation des ressources attentionnelles et enfin, (4) la confiance dans l'automatisation de la conduite. Au regard de ces résultats, nous en concluons que l'interface multimodale a permis d'améliorer les connaissances des conducteurs sur l'état des systèmes automatisés. Elle semble en revanche avoir un impact limité sur l'amélioration des compétences dans l'interaction des conducteurs avec l'automatisation de la conduite.

Globalement, notre thèse montre que les interfaces multimodales favorisent l'orientation de l'attention vers l'état des systèmes automatisés. Ces interfaces permettent d'améliorer l'interaction avec l'automatisation en améliorant la conscience des modes sans distraire le conducteur de sa tâche de conduite. Les méthodes mises en place pour concevoir et évaluer des interfaces multimodales permettent de s'assurer qu'elles orientent efficacement l'attention vers l'état des systèmes automatisés. Ces méthodes peuvent être exploitées dans d'autres domaines, tel que l'aviation, où humains collaborent avec des automates. Enfin, d'un point de vue industriel, notre thèse montre que les informations transmises par des interfaces de systèmes automatisés interfacés devraient être réparties sur plusieurs modalités sensorielles.

ABSTRACT

Partially automated systems are capable of temporarily replacing the human in the lateral and longitudinal control of the vehicle. However, they can return control of the vehicle to the driver when their operating limits are reached. It is therefore necessary for drivers to shift their attention from monitoring the activity of automated systems to controlling the vehicle. They need to identify the operating state of these systems at all times to avoid confusion. Multimodal and reliability interfaces seem to provide appropriate solutions to these problems. The current literature is not sufficient to conclude on the effectiveness of these interfaces in inducing efficient state awareness of automated systems. This thesis proposes to design and evaluate the effectiveness of interfaces exploiting sound, haptic, and peripheral vision modalities to orient drivers' attention on the operating state of assistance systems. A first experimental study carried out on the road showed that a multimodal interface with information oriented on the drivers' needs allowed a better understanding of the vehicle's functioning than a strictly visual interface with information oriented on the vehicle's state. Both interfaces caused mode confusion, and neither indicated the reliability of the automated systems.

Two experimental studies were then conducted to design and evaluate a peripherally perceptible limit approach indicator for automated systems. This indicator was considered to be useful information and was used to anticipate situations where automated systems suspended. A second group of experimental studies was conducted with the objective of verifying the capacity of audio and haptic interfaces integrated into the steering wheel to be perceived and comprehended by drivers. These were found to be effective in indicating the state of automated systems. The various interface elements tested previously were assembled in a simulator to evaluate their combined effect. In a longitudinal study, drivers were repeatedly and randomly confronted with four situations of recovery from the suspension of automated systems (sharp bends, traffic jams, fog areas, erased road markings). They had either a multimodal interface indicating the limits of the automated systems or a conventional visual interface. Awareness of the state of the automated systems, distribution of attention, and confidence in these systems were assessed. The results show a positive effect of the multimodal boundary interface on (1) the understanding of the functioning of the automated systems, (2) the driving performance during the suspension of the automated systems, (3) the orientation of attentional resources and finally, (4) the confidence in the automated driving. Based on these results, we conclude that the multimodal interface improved drivers' knowledge of the state of the automated systems.

However, it seems to have a limited impact on the improvement of drivers' skills in interacting with the driving automation.

Overall, our thesis shows that multimodal interfaces promote the orientation of attention toward the state of automated systems. These interfaces improve interaction with automation by enhancing mode awareness without distracting the driver from the driving task. The methods used to design and evaluate multimodal interfaces ensure that they effectively direct attention to the state of the automated systems. These methods can be exploited in other domains, such as aviation, where humans collaborate with automation. Finally, from an industrial point of view, our thesis shows that the information transmitted by interfaces of interfaced automated systems should be distributed over several sensory modalities.

CONTENTS

LIST OF FIGURES

List of Tables

LIST OF ACRONYMS

ACC Adaptive Cruise Control

ADAS Advanced Driver-Assistance Systems

AOI Area Of Interest

CLLE Cognition Langue Langage et Ergonomie

CMI Cockpit Multimodal Interactive

CR Correct Rejection

DCL Deviation from Central Lane

FA False Alarm

HTJA Highway and Traffic Jam Assist

HUD Head-Up Display

IPLA Indicator of Proximity to the Limits of Automation

LCA Lane Centering Assist

LDW Lane Departure Warning

MILA Multimodal Interface with indicator of Limits of Automation

RTLX Raw Task Load indeX

SDT Signal Detection Theory

SMI SensoMotoric Instruments

SRK Skill Rule Knowledge

TH Time Headway

TTC Time To Collision

VBI Visual Basic Interface

INTRODUCTION

Autonomous cars have been part of the cinematic landscape since 1968 in The Love Bug. While this early representation anthropomorphised the legendary Volkswagen Beetle, it has given way to purely mechanical and cold representations in the movies Minority Report, iRobot, and more recently in the TV show Westworld. In these representations, the vehicles are self-driving cars, autonomous and independent of humans' will. Even though the vehicles represented in these fictions are not yet reality, the number of on-board technologies in cars is exponentially growing. The automobile domain is now seeing a similar evolution as the aviation domain, where automated systems replace the human is some parts of the activity. These systems are useful for relieving pilots or drivers of cognitive loads. That is when the human operators are aware of what the automated systems do and when they have control over the vehicle.

Mont Saint Odile's crash in 1992 caused the death of 87 passengers (BEA, 1992). One of the reasons for that crash was mode confusions due to interface design (Sarter & Woods, 1995). Humans using modern cars now also have to collaborate with automated systems. Similarly to the aviation domain, incidents relative to automated systems begin to occur. In 2017, a fatal crash involving a vehicle driving in *Autopilot* mode caused the death of its driver. The vehicle collided with a tractor trailer crossing a highway in Florida. Neither the automated systems nor the driver brake or steer to avoid a collision. The Autopilot was not designed to respond to emergency braking situations related to crossing vehicles (NHTSA, 2017). It therefore appeared that most plausible explanation of the crash was a misunderstanding of the automation's limitation and functioning by the driver. The driver and the automated systems failed to collaborate to avoid a life-threatening situation. A first central question arises: how to characterise a safe collaboration between an automated vehicle and the Human operator?

An efficient collaboration starts with an understanding of the automation's functioning by the drivers. They should be aware of the operating state of automated systems and understand automation's limits. This phenomenon will be central in this thesis and will be referred to as mode awareness. A review of the existing literature on mode awareness and the factors impacting it will be presented. From a theoretical point of view, the goal of this thesis will be to contribute to the understanding of this phenomenon by making it possible to characterise it, assess it, and positively impact it. The study of existing vehicles, especially the one involved in

Florida's crash, will make it possible to identify how mode awareness is affected in current vehicles.

Then, a second central question arises: how can the two collaborators, the human and the machine, cooperate when the very nature of their functioning is different? The most viable solution probably resides in communication. The human-machine communication relies on interfaces. Selective attention, thanks to which humans focus their cognitive systems on the processing of a particular information, needs to be focused on the interface to extract the meaning of the message transmitted by the machine. If the machines that equip automated vehicles had all the means possible to communicate information to the operator, what would be the most effective and acceptable ones? Which would be the ones that would capture humans' attention the best, be understood, and would give them the most adapted trust in automation's capacities?

With the perspective of answering this question, a methodology of design and evaluation of new interfaces will be developed. This methodology relies on a systematic evaluation of each new interface. The goal is to estimate to what degree the interfaces orient attention toward relevant information to induce mode awareness. The evaluation process will be done with scenario-based studies, controlled laboratory tasks, simulator studies, and on-road studies. It will rely on the assessment of attitudes, behaviours, knowledge, visual fixations, rating scales, and on the collection of verbatim and interviews. It is only when all the functionalities are rigorously tested alone that they will be merged into a complete interface. The last central question is how do drivers learn to cooperate with automated systems depending on the information that interfaces present? A longitudinal study will be presented in which the effect of the new interfaces is assessed before and after multiple usage of automation. The answer to the previously mentioned questions will lead to theoretical implications about mode awareness and attention allocation. It will lead to design implications on the matter of information presentation. Finally, it will lead to methodology implications on mode awareness assessment. This thesis, addressing psychology and ergonomics challenges, took place in the context of a multidisciplinary project, making possible the design and development of new interfaces.

Cockpit Multimodal Interactive (CMI) Project was born of the collaboration of multiple industrial groups (Renault, Valeo, Arkamys, Saint-Gobain) and public research institutes (IRT SystemX, Bordeaux INP, CentraleSupélec). The collaborators pooled human resources ranging from human factor specialists, acoustics specialists, to engineers, developers, PhD students, and researchers. IRT SystemX hosted the project for the 4 years it lasted (2018 – 2022). The point of gathering such a wide range of collaborators was to share the expertise of each institution. This enabled the development, from scratch, of a driving simulator integrating automated systems, which were themselves synchronised with multimodal interfaces and, aiming to address ergonomic issues. This thesis is the result of a collaborative work between Renault Group and Cognition, Langage, Langue et Ergonomie (CLLE) Laboratory and covered psychological and ergonomic aspects of the interaction with automation.

The general goal of CMI Project was to build a multimodal and interactive cockpit for multiple degree of automated driving system, ranging from manual driving to highly automated driving. The objective was to study configurations of interfaces using multiple sensory modalities, depending on the degree of automation and on the profile of the user, to find the most efficient and acceptable way of transmitting information. A driving simulator was developed with a wide variety of interface: LEDs on the dashboard to indicate state of automation and direction of dangers, LEDs on the steering wheel to indicate the state of automation, LEDs encrusted in the windscreen to indicate the time remaining with automation, an instrument's cluster to inform on the vehicle's functioning, localized sound for alerts, sound in the headrest for information, haptic feedback in the steering wheel to indicate automation's state, and haptic feedback in the seat to alert on misuses of automation.

A framework composed of four subjects was developed to cover the most important aspects of the interaction with automation. Three of them focused on yet to come technologies, involving highly automated driving systems and innovative interfaces. The fourth subject, entitled IPLA and in which this thesis is included, covered existing technologies with shorter term application objectives. More precisely, Renault was interested in studying solutions to challenges occurring in commercially available partially automated vehicles. These challenges could engage the safety of drivers, as mentioned earlier, but also their comfort and therefore the acceptability of automated systems. CMI Project constituted a research framework in which Renault was able to experiment on interface solutions to address automation's challenges. Partial automation was the object of this thesis with the aim to respond to concrete challenges encountered by Renault's users such as how to understand limits and state of automation. For that purpose, CMI Project produced interfaces that could be developed by Renault with technologies already integrated in vehicles.

The purpose of this thesis is firstly to contribute to the understanding of key factors on the effect of interface design on the cooperation with automation. It aims to bring theoretical and methodological contributions to the scientific literature. Its results will be used by the collaborators of CMI Project and contribute to Renault's ongoing projects related to the interaction with automated systems. The results described in this manuscript will enable Renault to make choices regarding the integration of new interfaces in their future vehicles. More generally, it will contribute to expanding knowledge about the cooperation between humans and automation.

THEORETICAL SECTION

Automated driving systems can take multiple forms, each with their own specificities. The first goal of this section is to define automated driving, to understand its different aspects, and to apprehend the particularities of partial automation. In the first chapter of this section, we will find that partial automation raises challenges because its functioning can be difficult to assimilate for the human cognitive system. Among these challenges, we can find the orientation of attention toward relevant elements, mode awareness and adequate calibration of trust in automation. Interfaces, communication media between humans and automated systems, have been proposed by previous authors as a response to these challenges. A second goal of this section is to clarify how interfaces can address these challenges. A first interface solution highlighted in the literature is information of reliability of automation. The principles and some design examples of reliability interfaces will be presented. A second interface solution resides in interfaces exploiting multiple sensory channels, referred to as multimodal interfaces. The third goal of this section is to conduct a literature review on the capacity of multimodal interfaces to indicate the state of automated systems. In the second chapter of this section, a systematic review and a meta-analysis will be presented. The effect of interface modality on mode awareness will be presented. The literature about mode awareness is in the rise, but at the moment provide limited evidence. This leads to the fourth and last goal of this section: determining the directions of work explored during this thesis to extend the literature on the subject. The theoretical reasoning of this thesis will be developed in the third chapter of this section. The methodology followed for the design and evaluation of different interfaces will be explained. Finally, the plan of the thesis will be described.

CHAPTER 1 – AUTOMATION CHALLENGES AND SOLUTIONS

This chapter explains the principles of automated driving and the challenges it raises, particularly with partially automated driving. Among the main challenges that are cited are attentional resource allocation, its consequences on the identification of automated systems' state, and trust toward the automation. After presenting those challenges, the interfaces' role will be exposed to converge toward two promising types of interfaces: reliability interfaces and multimodal interfaces.

1. The Challenges of Automated Driving

Humans have used automation for several purposes over the course of history, such as safety, comfort, or skill enhancement (Janssen et al., 2019). Basically, any activity that can be more comfortable or safer thanks to technology, human will seek to automate it. Driving is no exception. The scientific literature identifies inattention while driving as the main cause of road accidents, given that more than 70% of car crashes involve inattention or fatigue (Dingus et al., 2006; Lemercier et al., 2014; Victor et al., 2015; Wang et al., 1996). In a society where the priority is to significantly increase road safety, automated cars appear to be a solution for reducing the human factor in road accidents. Before reaching automated vehicles that will replace the driver and reduce human factors involved in driving a car, vehicles are equipped with automated systems that gradually take in charge more and more parts of the driving activity. The interaction with such vehicles raises questions, especially after unfortunate events such as Florida's crash in May 2016 involving Tesla's Autopilot mentioned in Preamble. According to the National Transportation Safety Board's report (2017), this crash was caused by misuse of the automated systems by the driver, who was not aware of how the systems worked and what their limitations were. One explanation is that automated systems of vehicles are mostly black boxes. Drivers cannot directly see and understand how the automation is going to react in a situation. Therefore, it is difficult to act adequately and rely on the system in the appropriate way. However, the black box of automated vehicles possesses windows: the interfaces. An interface is understood here as any communication between the driver and the vehicle, whether it takes place via auditory, haptic and visual displays, as well as via controls. The interface constitutes the bond between the driver and the vehicle. Solutions to improve the interaction with automated driving systems might therefore reside in it. Before addressing the challenges of automation, it is necessary to understand how can automated system be characterised.

1.1. From Manual Driving to Autonomous Vehicles

Totally automated vehicles should be commercially available around 2050 (SAE International 2018). Before that, different levels of driving automation will become available (ERTRAC Task Force 2015). Six levels of automation have been established by the Society of Automotive Engineers to classify the roles of automated systems and humans depending on the degree of automation (SAE International 2018). According to Michon's (1985) classic model, driving requires three stages of control: strategic, navigational, and operational. *Strategic control* refers to the control of the overall direction (e.g., traveling from A to B), *navigational control* refers to immediate decisions (e.g., take the next turn), and *operational control* refers to the effective lateral and longitudinal control of the car. Depending on the level of automation, different stages of control are taken in charge by automation. Level 0 corresponds to manual driving. Humans take in charge all three stages of control. At Level 1, the system only takes charge of either longitudinal control or lateral control. Most vehicles equipped with Level 1 systems take charge of longitudinal control, meaning that it can automatically maintain either a target speed or the same speed as the vehicle in front (Adaptive Cruise Control, ACC). Level-2 vehicles are referred to as partially automated vehicles, as they feature automated systems that can partially perform the driving activity. Two Advanced Driver-Assistance Systems (ADAS) take in charge of both longitudinal (ACC) and lateral (Lane Centering Assist, LCA) movements. The term automated systems, or automation, will be used in this manuscript to encompass the ADASs that take the driving activity in charge. Level-3, Level-4 and Level-5 vehicles are classified as highly automated, as the driver can delegate driving to the system and engage in secondary tasks. At these levels, the vehicle respectively takes in charge all three stages of vehicle control.

1.2. The Particular Case of Partially Automated Vehicles

A confusion often occurs between autonomous cars and automated cars. Autonomous systems can be defined as the following:

« autonomous systems are generative and learn, evolve and permanently change their functional capacities as a result of the input of operational and contextual information. Their actions necessarily become more indeterminate across time. » (Hancock, 2019, p. 481)

While automation can be defined as the following:

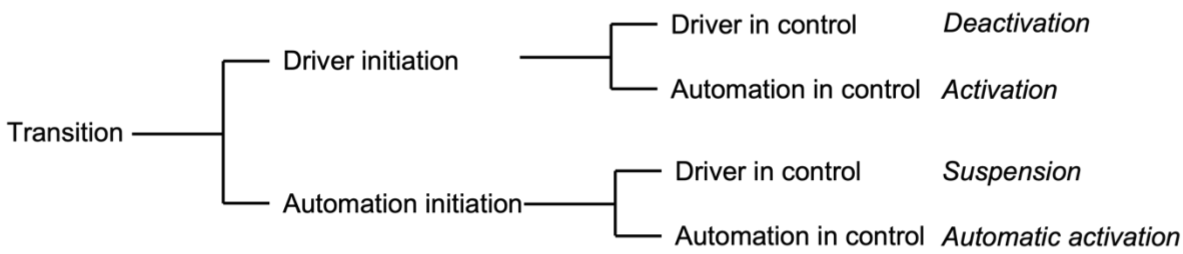
« automation is defined as 'automated systems are those designed to accomplish a specific set of largely deterministic steps, most often in a repeating pattern, in order to achieve one of an envisaged and limited set of pre-defined outcomes. » (Hancock, 2019, p. 481)

The autonomous vehicles that are often referenced in the media are in fact Level-5 vehicles. In such vehicle, the humans are not implicated in the driving task. If an incident occurs, human passengers cannot control the vehicle, and therefore be held responsible. It is different in Level-2 vehicles, also referred to as partially automated vehicles, as both the driver and the automated systems act on longitudinal and lateral controls of the car. More importantly, in such vehicles, the human operator is the only one responsible for supervising the driving activity. If any problem occurs, the driver is to blame. Automated systems are mere executants of accelerating, braking, and turning actions. This thesis dissertation is focused on Level-2 vehicles, which are the most advanced automated systems commercially available today. Renault works on these vehicles and seeks to improve the interaction with it. CMI project aimed to respond to the challenges posed by automation with the use of multimodal and interactive interface. It is first necessary to identify these challenges.

In partially automated vehicles, frequent transitions of control of the vehicle occur between the driver and the automated systems. Lu & de Winter (2015) highlighted the different types of transition of authority over the vehicle depending on which antagonist, the drivers or automation, initiate the transition and which antagonist assume the control of the vehicle after the transition. Transitions can be initiated by the drivers or the automation. After the transition, either the automation or the drivers control the vehicle (see Figure 1 for an illustration). To illustrate the differences in the transition of control, we will explain a particularity of partially automated vehicles which is that automated systems need specific conditions to operate. Along these conditions, we can find having clear lane markings, not reach a certain threshold of lateral and longitudinal acceleration, and not having the sensors blurred, just to name a few. When these conditions are complete, the drivers can turn on automated systems. A driver-initiated transition of control occurs from the driver to the automation (i.e., *activation*). While these conditions are met, automated systems can work indefinitely. However, if only one of these conditions is not met anymore, automated systems will cease functioning suddenly. At this moment, the control of the vehicle transits from the automation to the driver at the initiation of the automation, because of non-completion of the system's conditions. This phenomenon will be referred to in the rest of the manuscript as a *suspension* of automation. This kind of transition is different from a *deactivation*, in which the control of the vehicle is given to the drivers, at the initiation of the drivers themselves (e.g., pressing the "Off" button). For the drivers, suspensions are more complex than deactivations, because they are not always prepared to recover control of the vehicle (Lu & de Winter, 2015). In some situations, automation suspends for a brief period (e.g., when road markings are erased) and automatically activate when correct conditions are recovered. The automation initiates the transition of control from the driver to automated systems (i.e., *automatic activation*).

Figure 1

Categorisation tree of transition of control depending on the initiator, inspired by Lu & de Winter (2015).



To add complexity, the automation of these vehicles is composed of two automated systems (i.e., ACC and LCA). A system can reach its limits and not the other one, leading to a transition of control of only one dimension of operational control (e.g., only lateral control and longitudinal control). Some suspensions of automation are caused by a reached limit that cannot be anticipated (e.g., defective sensors). Other suspensions are caused by a gradual degradation of conditions, meaning that an estimation of the risk to reach the limit can be predicted. Situations in which automation gradually reach their limits were selected with experts of Renault group because an anticipation to reach limit could be calculated, and the interface could represent this prediction. These situations are presented in Clio 5 user manual[1] as situations in which automated systems should not be used or used carefully. Four situations were selected:

[1] see https://fr.e-guide.renault.com/fra/Clio-5/Assistant-Autoroute-et-Trafic, retrieved on May, 12th, 2020.

bent roads, slow traffic jams, fog areas and erased road markings. During a trip to vacations on the other side of the country, drivers are likely to encounter such situations. Along the road, the drivers will face different challenges and use different parts of their cognitive system to interact with automation. One of the central challenges that drivers will face are efficient attentional resource distribution.

1.2.1. Attentional Resources

A first challenge that comes across drivers when using partially automated vehicles is to correctly distribute their attentional resources. Even through driving engages the perception and processing of multiple sensory modalities, visual information is the main one (Sivak, 1996). Conscious control of information processing involves attentional resources. The attentional resources are limited and selective attention can be allocated to one element at a time (Kahneman, 1973). While driving with partially automated systems, two tasks are usually performed: monitoring the correct activity of the automated systems and controlling the vehicle when necessary (Carsten & Martens, 2019). The attentional resources are distributed to either one of these two tasks. For example, when using LCA on a straight highway with low surrounding traffic, the drivers mainly allocate their attentional resources to monitoring the correct functioning of automation, monitoring the behaviour of other vehicles, listen to music or talks with passengers. When reaching a limit of automated systems, because of erased road markings for example, the drivers' attentional resources will be allocated to controlling the vehicle. The driver must switch the focus of attention between the two tasks: monitoring and controlling. However, humans can have difficulty maintaining sustained visual attention toward a source of information in which little happens (Bainbridge, 1983). Therefore, prolonged periods of automation usage can lead to "out of the loop phenomenon", situations in which drivers fail to detect errors and to react appropriately to automation failures (Endsley & Kiris, 1995). The driver is not engaged in the driving task and fail to switch the locus of attention from monitoring to control. The out-of-the-loop phenomenon in automated systems can be explained by poor situational awareness.

Situational awareness can be defined as:

« the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future » (Endsley, 1988, p. 97)

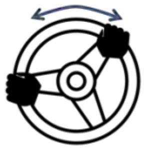
On a highway, this faculty allows the drivers to perceive the surrounding cars, comprehend their behaviours, and project their future trajectory. While automated systems are activated, the drivers are mainly concerned by monitoring the activity of automation. They shift from an active role to a passive role, which reduces their vigilance and lead to a reduction of situational awareness (Endsley & Kiris, 1995). In situations in which automation suspends suddenly, for example when passing a bend road with too elevated speed, the attentional resources must be rapidly allocated to controlling the vehicle. The drivers need to comprehend the new mode of automation and that a takeover of the direction of the car is required. If drivers fail to identify that a transition from automated to manual steering is happening, they can face mode confusions, a false estimation of the mode of automation (Baltzer et al., 2017). The vehicle will continue its trajectory and in the worst-case scenario, it will leave the road.

1.2.2. Mode Awareness

When a transition of control from the vehicle to the driver occurs, drivers must understand that the automated system has been suspended and that they must take over the vehicle's control. This type of suspension can come as a surprise, as the human and the vehicle perceive elements of their environment differently. Several surprising suspensions were experienced by Endsley (2017) during a 6-month naturalistic driving study in a Tesla Model S. For example, the automated system unexpectedly shutdown in a bend, probably because the lateral acceleration was too elevated for it. This resulted in *Automation surprises,* situations in which the system's behaviour is different from the user's expectation (Sarter et al., 1997). A driver in manual driving would have passed the same bend. In this case, the capacities of action of automation are inferior to that of humans. The most common automation surprises a user can encounter are the result of mode confusions. *Mode confusion* refers to a situation in which a user thinks a system is in a different mode than is actually the case (Baltzer et al., 2017). For example, a driver may enter a bent road believing the LCA is activated when in fact it is not. When a mode confusion occurs, the user may commit a *mode error*, involving the execution of an intention that is adapted to one mode when in fact the system is operating in a different mode (Sarter et al., 1997). For example, the driver may not turn the steering wheel in a curve, owing to a mistaken belief that the LCA is activated (see Table 1).

Table 1

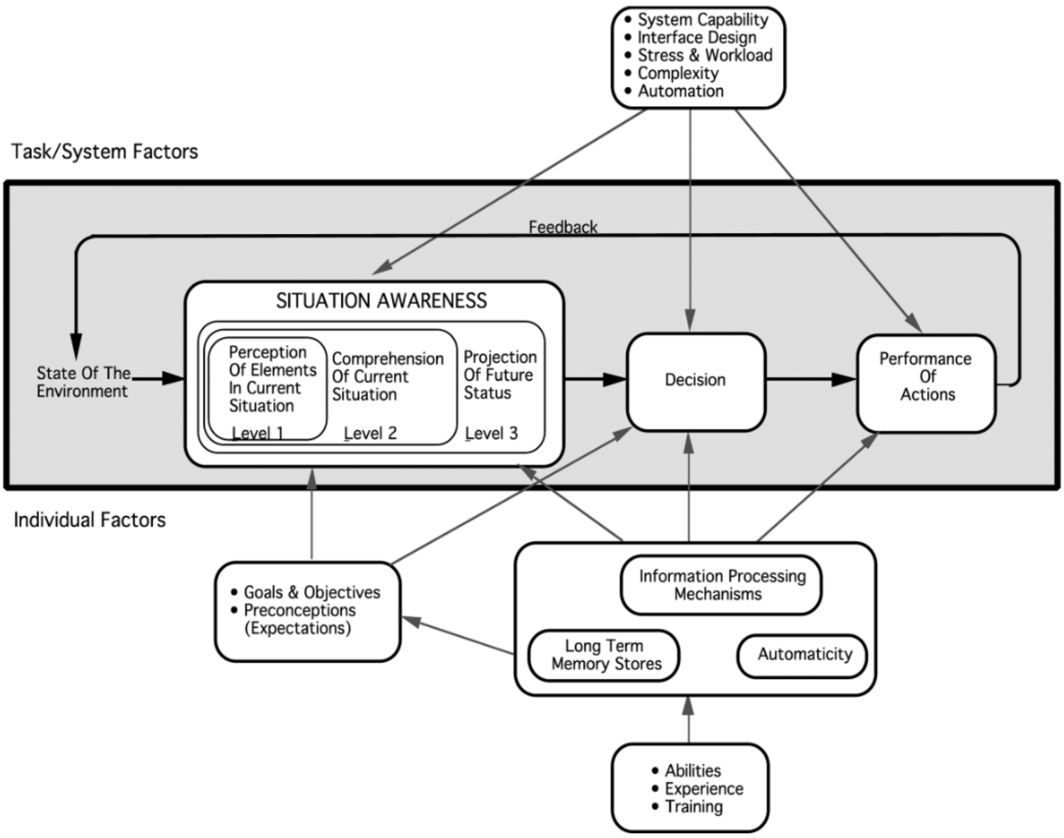
Description of mode awareness depending on automation's state and drivers' behaviours

Status of automated system	Vehicle's behaviour	Driver's behaviour	Mode awareness
LCA stays activated			*Accurate mode awareness*
LCA suspends			*Mode error Poor mode awareness*
LCA suspends			*Accurate mode awareness*

Mode confusions and mode errors are related to a deficient mode awareness (Kurpiers et al., 2020). *Mode awareness* refers to both an understanding of how automated systems work, which relies on mental models (i.e., drivers' representations of how these systems operate), and an awareness of these systems' state of activation (Monk, 1986). Mode awareness is a subcategory of situational awareness, in that it shares the same features, namely perception, comprehension and projection (Endsley, 1995; see Figure 2). The perception level corresponds to the differentiation and identification of the elements of the signals emitted by the interface. The comprehension level refers to the correct association between the signals and their meaning, considering the operator's goals. The projection level refers to the projection of a probable outcome of a future event, depending on the perception and comprehension of information available. Individual factors and task/system factors can impact mode awareness. In individual factors, experience (i.e., time spent using the system) and training (i.e., preparation to use the system, see Endsley & Garland, 2000) influence long-term memory, automaticity (i.e., processing without attention), and information processing mechanisms, which influences goals and objectives, and directly impacts mode awareness. Task and system factors are the system's capability, the interface design, the stress and workload provoked by the task, its complexity, and the issues related to automation interaction (see Figure 2). This description will focus on interface design and workload as system factors and on training and experience as individual factors.

Figure 2

Situational/mode awareness model, inspired by Endsley (1995).

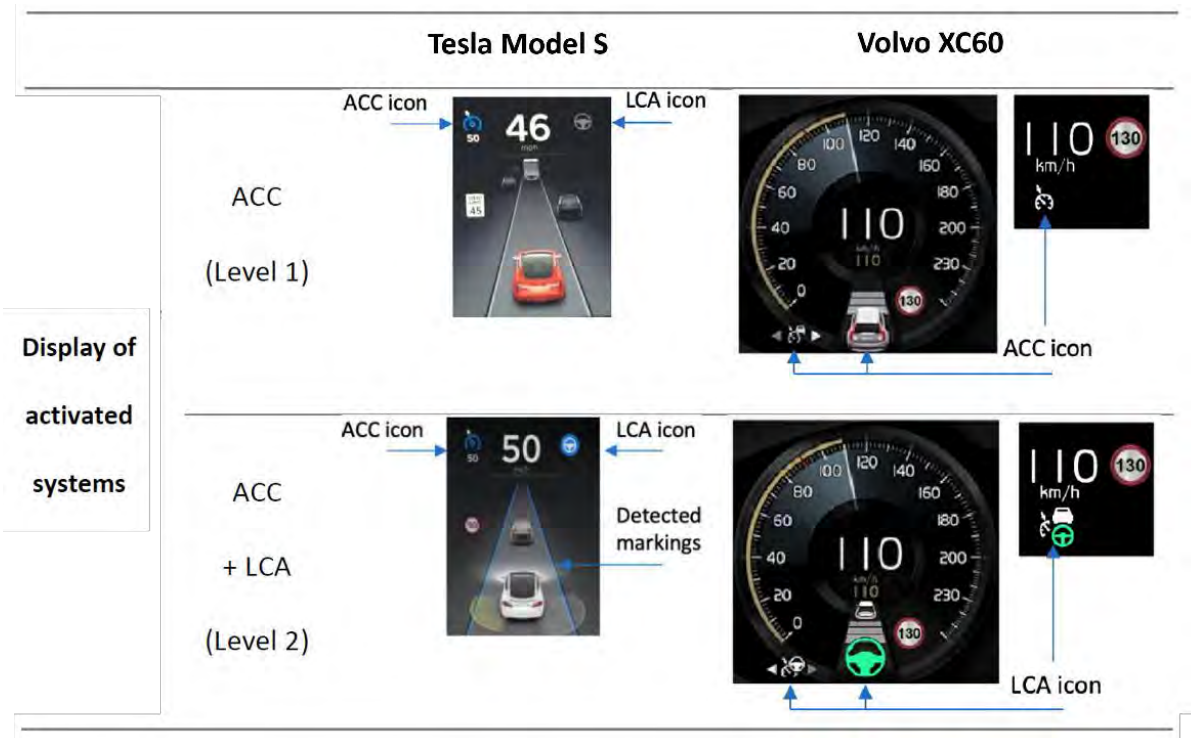


Workload and mode awareness

Instruments' cluster of Level-2 vehicles usually indicates the state of automated systems (see Carsten & Martens, 2019 for examples). The current mode of automation is usually continuously displayed with icons on the instrument's cluster. In some vehicles, different icons in different locations are used to represent the activated states of ACC and LCA (see Figure 3 for an example of the Volvo AXC60). In other vehicles, the same icon changes its form or color to represent the ACC or the LCA (see Fig. 3 for an example of the Tesla Model S).

Figure 3

Representations of automated systems visual icons of Tesla Model S and Volvo XC60, adapted from Monsaingeon et al. (2021).



The driver gauges the state of the automated systems by identifying the meaning of these icons and ignoring any irrelevant and distracting icon (i.e., noise). As the number of interactive technologies in vehicles increases (Regan et al., 2009), the number of visual icons on the instrument's cluster also increases. The drivers therefore have to ignore more and more distracting noises before finding the icon they are looking for. The difficulty is that the drivers might have to do this in time critical situation. For example, if automated systems suspend in a bend, many visual information has to be processed on the road to correct the trajectory of the vehicle. At the same time, the drivers must gaze at the cluster to identify which AUTOMATED SYSTEMS suspended. As stated earlier, the attentional resources being limited, a more important workload will occur, reducing mode awareness. Some interfaces might reduce workload by distributing the demands of the interface on other sensory modalities than the visual one, resulting in faster takeovers (Zhang et al., 2019). However, the drivers might identify situations in which automation suspend with training.

Knowledge About Automation: Mental Models

Experience with automated system plays an important role in correct understanding of the role one has to play in the interaction (Solís-Marcos et al., 2018). Interactions with automated driving systems allow to forge a representation of its purpose, form, functioning, state and structure, which can be merged into the term *mental model* (Seppelt & Victor, 2020). Driving-related mental models are influenced by three main factors: experience, training, and interface transparency (Endsley, 2017). Experience came into play in the study of Forster et al. (2019), as the more situations the participants encountered, the more accurate their mental models became. These findings were supported by the study of Strand et al. (2018), where three purchasers of Level 2 vehicles were interviewed as they familiarized themselves with the system. Results revealed that although they refined their mental models, their primary representations of the system influenced the subsequent formation of their mental models. Blömacher et al. (2020) reached the same conclusion. They gave correct or incorrect information to drivers regarding the functioning of the system before they used a Level 3 vehicle. Results revealed that the veracity of the information given at the beginning influenced the formation of mental models and reaction time at takeovers. Mental models can also be influenced by the interface's information. In a study on Level-1 vehicles, Seppelt and Lee (2019) concluded that the development of mental models is tied to the quality and form of information transmitted by the interface. Experience, training, and interface design are factors to take into account to evaluate mode awareness, which can be done with different techniques.

Evaluation of Mode Awareness

As proposed by Kurpiers et al. (2020), the assessment of mode awareness can be performed by measuring three dimensions: the driving behaviour of the drivers, their eye behaviour, and their mental models. The behaviour of the drivers should be adapted to the mode of automation. After a system-initiated transition of control to the drivers, the deviation from central lane or Time To Collision to the followed vehicle (TTC) should reveal that the drivers are in control of the vehicle. The ocular behaviour of the drivers should reveal that their gaze is fixed on the exterior of the environment when a takeover occurs. When in highly automated driving, drivers can gaze at the interior, but when switching to partially or manual driving, the gaze should be located to the exterior environment. The mental models, evaluated through questionnaires regarding the functioning of automation depending on the situation, should be accurate. Evaluations of gaze proposed by Kurpiers et al. (2020) are designed to Level-3 vehicles and is based on a comparison of gaze duration on the exterior environment between Level-2 and Level-3 driving periods. It is therefore important to assess how is evaluated mode awareness in vehicles equipped with only Level-2 automated systems. This question will be answered in the systematic review of Chapter 2. Mode awareness and trust share an intricate relation. As consequence of accurate mode awareness, trust toward automation should be appropriate to the situation, because the drivers are aware of their roles and are aware of the limits of automation.

1.2.3. Trust Calibration

Trust can be defined as the attitude an operator has toward an agent that helps him/her achieving a goal in a situation where uncertainty and vulnerability are involved (Lee & See, 2004). If automation does not accomplish the goals that it is meant to achieve, breakdown of trust can be observed and lead to decisions not to use it (Parasuraman & Riley, 1997). More precisely, failure of automated systems can lead to decreasing trust if the cause of the failure appears random or from internal sources (e.g., software; see Bisantz & Seong, 2001). On the other hand, if the operators place too much trust in the automated system, over trust can be observed leading to inappropriate use of automation. With highly automated vehicles, drivers over trusting the automated system gazed less at the road (de Winter et al., 2014) or failed to take over correctly when needed (F. O. Flemisch et al., 2014). Trust in automation is linked to the drivers' mental model of automated systems (Seppelt & Lee, 2019). If they believe that automation can deal with almost all types situations, they will place too much trust in it. Therefore, trust in automation should be appropriately calibrated to the automated system's capacity and limits. One way to calibrate trust is to indicate the reliability and limits of automation. Current interface rarely displays information regarding the automation's reliability, the most common one being the correct detection of lane markings (see Chapter 4 examples of representations). Yet, informing the drivers of the automated systems' capacity and reliability allow them to place adequate trust in it (Helldin et al., 2013). Interface design could also address the challenge of calibrated trust by indicating the limits of automation. An evaluation of utility and design of an interface indicating limits of automation and its relation to prior trust in automation will be exposed in Chapter 5. An evaluation of the effect of this interface on trust will be presented in Chapter 9. The issues of attention distribution, mode awareness, and trust calibration, are all influenced by interface design. We will now investigate how interface design can address these issues.

1.3. Interface Principles to Address the Challenges of Automation

Based on the considerations presented above, Carsten and Martens (2019) have identified six goals that need to be addressed to improve the interaction automated driving systems:

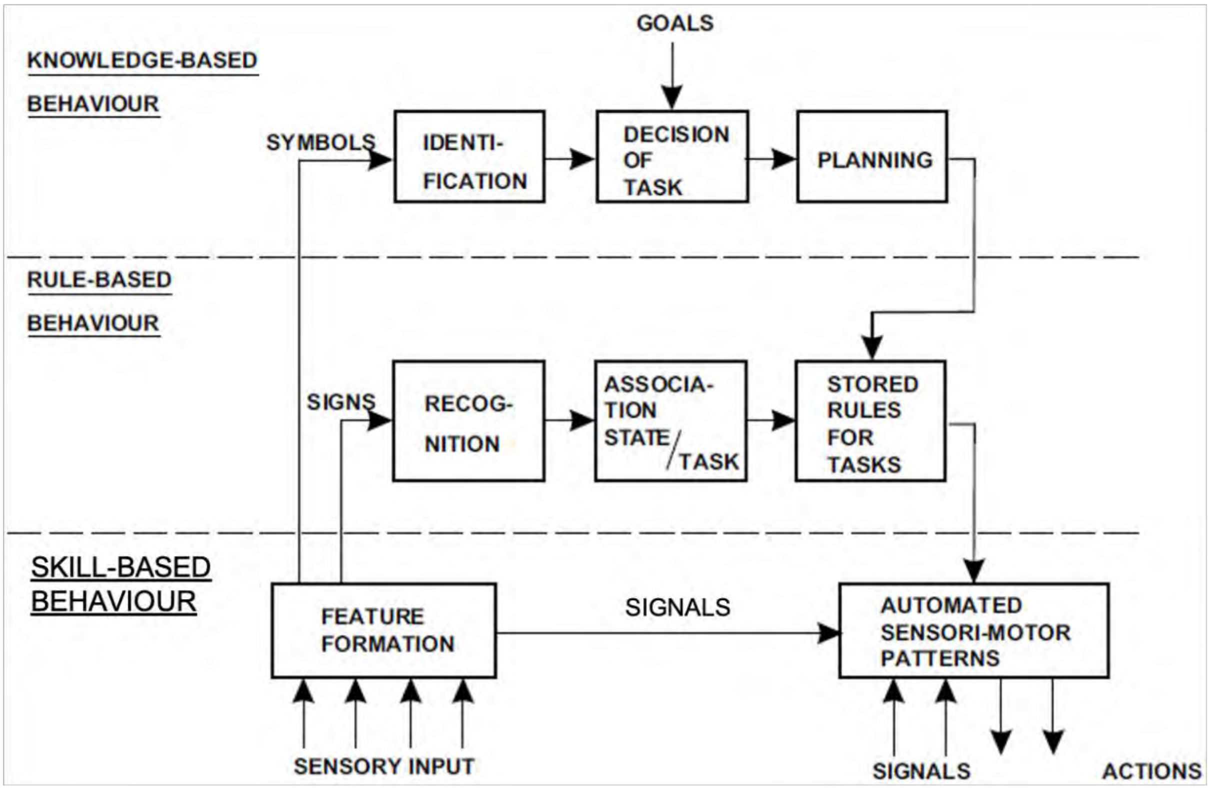
« (1) Provide required understanding of the automated vehicles capabilities and state (minimise mode errors); (2) Engender correct calibration of trust; (3) Stimulate appropriate level of attention and intervention; (4) Minimise automation surprises; (5) Provide comfort to the human user, i.e. reduce uncertainty and stress; (6) Be usable. » (Carsten et Martens, 2019, p. 5).

Challenges (1), provide required understanding of automated vehicles capabilities and state, and (4), minimize automation surprises, can be gathered under a common global of increasing mode awareness, as they both refer to the correct perception, comprehension, and projection of automation's mode (Endsley, 1995). As proposed by these authors, interface design can address these challenges. In this thesis, we evaluated to what extent interfaces address the following goals: Stimulate appropriate level of attention and intervention, increase mode awareness, engender correct calibration of trust, be usable. Those challenges were the only ones addressed in this manuscript, because the manipulated parameters of interface have been shown to impact the factors of these challenges only.

To address these issues thanks to interfaces, Carsten and Martens (2019) propose to favour at the maximum non-reflexive behaviours. This guideline is based on Rasmussen's Skill Rule Knowledge (SRK) model (1983) of performances of skilled human operators (see Figure 4). The Skill-based behaviour refers to sensory-motor behaviours that take place without conscious control. The Rule-based behaviours refers to behaviours such as procedures of familiar routines. The performance is oriented toward a goal that is often implicit and dictated by the situation releasing the rule. Both the Skill and Rule based behaviours are fed by signs. Signs are continuous information about the physical world. The Knowledge-based behaviour refers to behaviours that are goal oriented. The goals are explicit, and the human follows a plan that is dependent of the internal representation of the system and therefore to a higher conceptual level. Symbols feed knowledge behaviours in the form of explicit references to functional properties. Maintaining the operator at a Skill level should induce fast reaction time with low cognitive process. The following section will present the interface principles that were explored in this thesis.

Figure 4

Simplified version of the three levels of performance of skilled human operator, adapted from Rasmussen (1983).



2. Interface Design Solutions

The interface of the vehicle is the link that bounds the driver to the automated system. As mentioned earlier, interface providing information about automated systems impact attention allocation, mode awareness, and trust. We still must evaluate how interface design can impact the interaction, and if there are interface designs that are more efficient than others. The challenges of automation are present with vehicles that possess mainly visual interfaces without continuous information about limits of automation. We investigated two main axes of research of the interface design: reliability displays and multimodal interfaces. Reliability displays offer potential benefits on attentional resources distribution, mode awareness and calibration of trust. Multimodal interfaces present the advantage of communicating without disturbing the drivers' visual resources, offering benefits on attentional resources distribution and mode awareness.

2.1. Indicating Reliability of Automation

Automated systems of cars possess limits. In order for the drivers to understand their automated systems, they must build accurate mental models and calibrate their trust (Seppelt & Lee, 2019). A way to make drivers understand their automated systems is to display reliability of automation. By indicating the proximity to the limits of automation, drivers can judge the degree of reliability of automated systems in a context and act accordingly. It can help drivers anticipate suspensions of automated systems, improve takeover performances, improve mental models, and help calibrate trust (Beller et al., 2013; Helldin et al., 2013; Monsaingeon et al., 2021; Seppelt & Lee, 2019). This information is often communicated by reliability displays that can take the form of a simple icon (Beller et al., 2013), gradual displays in the form of a bar graph (Helldin et al., 2013), or continuous visual representation of proximity to limits of automation (Seppelt & Lee, 2007).

2.1.1. Effects of Reliability Displays

Indicating the reliability of automation, or the proximity to its limits, impacts attention, mode awareness, and trust. In a study where the attentional demand of aircraft piloting was mimicked, reliability information allowed the participants to better detect automation failures (Bagheri & Jamieson, 2004). Their attentional resources were more easily directed to the relevant information. These results were confirmed by the experiment of Wickens et al. (2000), in which predictive information about automation outcome helped aircraft pilots to allocate their attention and improving their performances. These results highlight the potential of reliability display to direct attention to relevant information and react faster to failures in aircraft. Similar effects should be observed in automobiles by helping the drivers to anticipate automation suspension. It should increase the understanding of drivers regarding automation. In a study by Seppelt and Lee (2019), continuous indication of limits of automation led drivers to better mental models of an ACC than a classical on/off display. Moreover, indicating the proximity to the limits of automated systems can allow to better understand when automation presents a risk of suspension and avoid these situations (Monsaingeon et al., 2021). These findings indicate the potential of an IPLA to positively impact mode awareness. With better attentional allocation and better understanding comes better performances. In a driving simulator study, Seppelt and Lee (2007) used an IPLA to inform on limits of an ACC which improved performances to anticipate the front vehicle's behaviour. Beller et al. (2013) observed similar effects. In partially automated vehicles, TTC was increased thanks to an IPLA. An IPLA can also improve takeover performances in situations of bad weather (Helldin et al., 2013). These authors also found that IPLAs allow to calibrate trust to the actual capacity of automated systems, or more generally to increase trust in automation (Beller et al., 2013). Seppelt and Lee (2019) proposed that mental models and trust are intimately related. By improving the drivers' understanding of the automation's limits, trust should be calibrated accordingly, which was confirmed by their results. Interfaces indicating the limits of automation should address several challenges of automated driving: stimulate appropriate attention, induce accurate mode awareness, and calibrate their trust accordingly. However, some questions remain. There are several limits of automation depending on the environmental conditions. How do drivers comply with IPLAs in this context? Do they comply solely with the information about the reliability of the automation, or do they judge the situation according to all the environmental conditions? Moreover, how do experienced drivers, who already understand the system's limitations, react? These questions will be addressed in Chapter 5. In the studies mentioned here, the design of the IPLAs were all different. We will now investigate the different designs of IPLAs to find the solution that better suits our objectives.

2.1.2. Design of Reliability Displays

Reliability and approach of the limits of automation have been represented in multiple ways (see Figure 5 for examples), either in two levels of colors or shapes (Beller et al., 2013; Wintersberger et al., 2019), or continuous representations (Helldin et al., 2013; Kunze et al., 2019; Seppelt & Lee, 2007). Even the simplest binary form (green color for "everything ok" and red color for "automation is likely to misbehave") has been proven to make automation safer (Wintersberger et al., 2019). Beller et al. (2013) studied a representation of automation reliability in their pioneering work on Level-2 simulated vehicles. In their study, drivers of partially autonomous vehicles were presented with two states of automation reliability in two weather conditions. The reliability display depicted a doubtful face when the automated system could not cope with the situation (fog). When the automated system was not reliable, it did not break when approaching another vehicle. Participants therefore had to manually brake. The minimum time to collision was found to be longer when the automation reliability information was displayed. In a simulation study by Helldin et al. (2013), participants drove Level-2 automated vehicles. Lateral and longitudinal control was undertaken by the vehicle on a mountain road with varying snowfall density. Participants in one of two groups could see a 7-

point reliability display. The denser the snow, the smaller the reliability representation. Continuous displays were proposed by other authors.

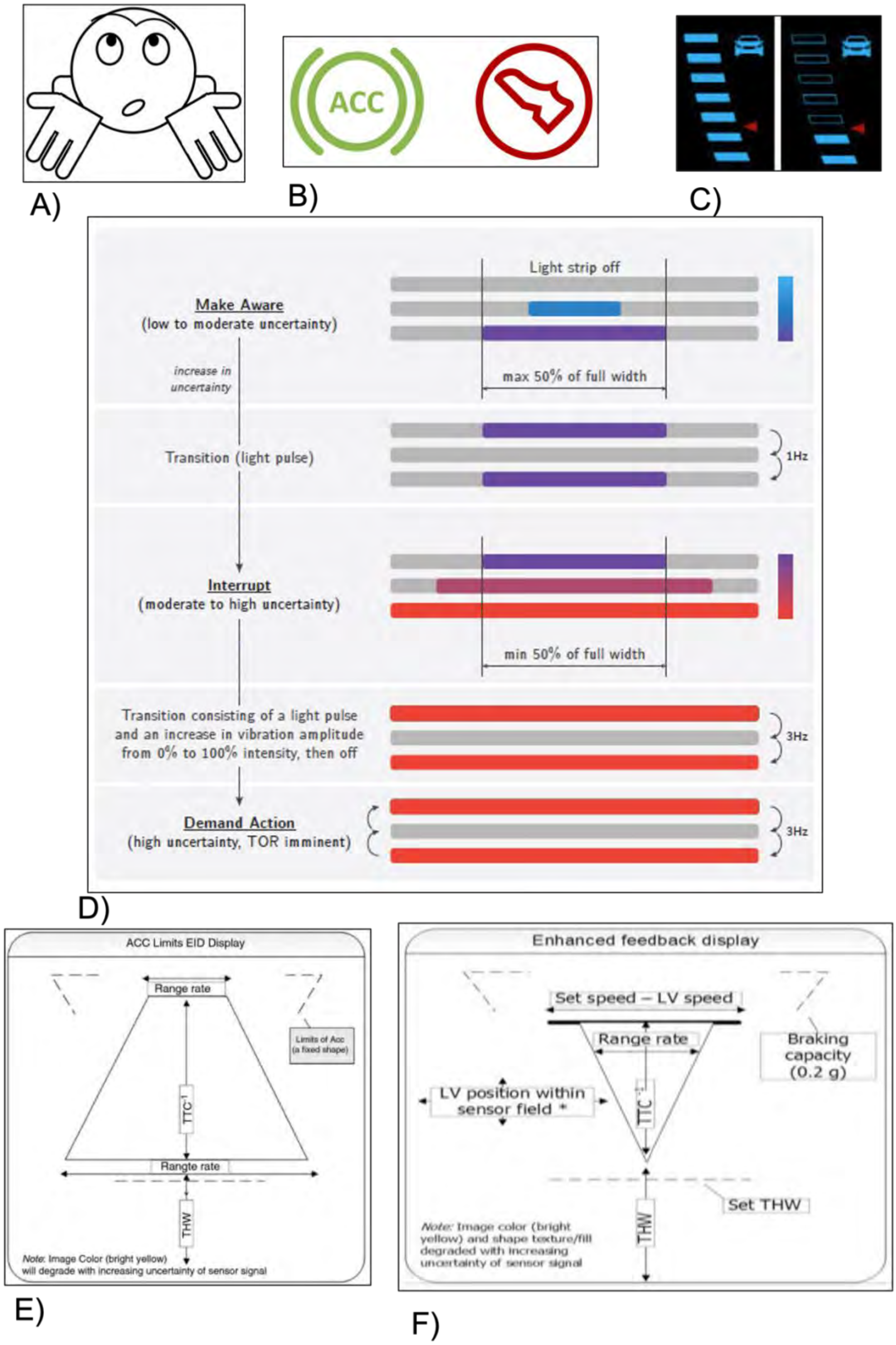
In Seppelt and Lee's (2007) pioneering research on how to represent the limits of automation, participants were placed in a simulator with Level-1 automation that relied on the ACC system. An IPLA was displayed. This ecologically designed interface featured a triangle that decreased in size when the systems approached the limits of automation. The representation of the limits of automation proved beneficial, particularly in urgent braking situations. This design was improved in a later study by integrating the position of the lead vehicle relatively to the drivers' vehicle (Seppelt & Lee, 2019). In addition to limits of braking capacity of ACC, this display represented the limits of detection of the lead vehicle, limits of sensors and setting exceedance. The triangle moved from left to right depending on the movements of the lead vehicle. The state of ACC was represented with shades of grey. All the cited representations of limits of automation require the drivers to gaze at the instrument's cluster to acquire information.

Recent studies proposed to display information on limits of automated systems in peripheral vision of the drivers, to avoid distraction of a visual task performed in focal vision (Kunze et al., 2019). On a LEDs bar, variations of color and pulses indicated the proximity to the limits of automation in a highly automated vehicle. This display allowed drivers to better distribute their gazes and increase their performances in takeover situations. Participants were able to perform secondary tasks, which would not be the case in Level-2 vehicles. Would a peripheral display be efficient in Level-2 vehicles, to inform drivers on risks of suspension of automation while they keep the road in focal vision? In all the above-mentioned studies, drivers appeared to take the reliability information into account, which impacted their driving performance, mental models, or trust. The question remains about how an IPLA in should be a partially automated vehicle, and will be addressed in Chapter 6. To complete the IPLA, interface using different sensory modalities than vision could convey information about the automation's state.

Figure 5

Examples of reliability representations extracted from the following articles.

A) Beller et al. (2013); B) Wintersberger et al. (2018); C) Helldin et al. (2013); D) Kunze et al. (2019) ; E) Seppelt & Lee (2007); E) Seppelt & Lee (2019).



2.2. Multimodal interfaces

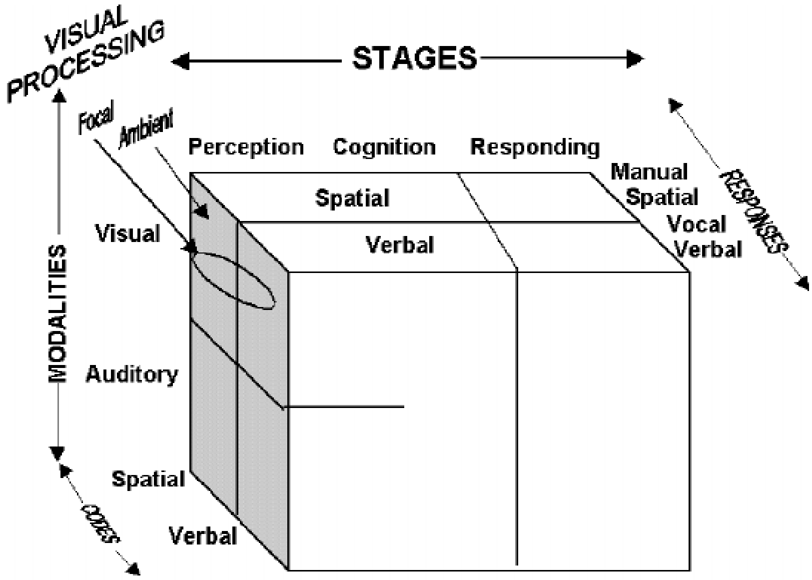
As stated earlier, attentional resources are limited. Humans might have one pool of attentional resources which can be distributed into different channels.

2.2.1. Multiple Resources Theory

Based on results of dual task experiments, Wickens proposed a multiple resources model (2008). This multidimensional model postulates that humans' attentional resources can be represented in different dimensions, some of which share resources and others do not. If two tasks share common resources, a decrease of performance is likely to be observed compared to if two tasks do not share resources. According to this model, there are two stages of information processing: perceptive/cognitive stage and response stage. There are two codes of processing: spatial and verbal codes. This is illustrated by the different processes involved in working memory while maintaining spatial information on the visual sketchpad and maintaining verbal information in the phonological loop (Baddeley, 1986). Finally, the perceptive possess two modalities of resources: visual and auditory perceptions. Visual perception can be cut into two resources: focal and ambient resources, referring to peripheral vision (see Figure 6). This model predicts that when dealing with a task that involves visual perceptions and processing in parallel with a task involving auditory perception and processing, humans would perform better than when dealing with two tasks involving visual perceptions and processing.

Figure 6

Representation of the multiple resources model, extracted from Wickens, 2008.



2.2.2. Potential effects of multimodal interfaces

Based on this statement, multimodal interfaces should allow to distribute the attentional demands of the interface on the multiple sensory channels, thereby reducing cognitive load compared to a situation where all demands are directed to a single sensory channel (Wickens, 2008). In a meta-analysis, Zhang et al. (2019) evaluated the effect of the modality of presentation of takeover requests on take over time in highly automated vehicles. The effect of visual takeover requests was compared with the effect of visual, auditory and vibro-tactile takeover requests. The results revealed that takeover time was significantly shorter when the interface was multimodal compared when it was only visual. These results concern highly automated vehicles with takeover requests. In partially automated vehicles, which are of interest to us here, there are no takeover requests. The suspension of automation is immediate. The question therefore remains about how multimodal signals would help drivers takeover in such vehicles. Based on results on highly automated vehicles, it should allow to answer the first challenge of partial automation, which is to allocate attention efficiently. The takeover requests studied in Zhang et al. (2019) had a simple meaning which was that a take over was necessary. Auditory and haptic interfaces can convey more complex information and potentially inform on the state of automated systems, which would increase mode awareness. The effect of multimodal interfaces, studied in the literature, on mode awareness will be investigated in Chapter 2.

Points clés

- Les systèmes de conduite partiellement automatisés peuvent prendre en charge le contrôle longitudinal et latéral du véhicule pendant que les conducteurs supervisent l'activité des systèmes automatisés.

- Les ressources attentionnelles des conducteurs sont principalement allouées à la surveillance de l'activité des systèmes automatisés ou au contrôle du véhicule.

- De faibles performances de reprise de contrôle du véhicule, des confusions de mode, et une confiance inadéquate ont été observées dans les véhicules partiellement automatisés.

- Il a été démontré que les interfaces indiquant la fiabilité de l'automatisation améliorent l'allocation de l'attention, les modèles mentaux et le calibrage de la confiance.

- Les interfaces multimodales ont un effet bénéfique sur l'allocation de l'attention et potentiellement sur la conscience des modes.

Key points

- Partially automated driving systems can take in charge longitudinal and lateral control of the vehicle while the drivers supervise the activity automation.

- The drivers' attentional resources are mainly allocated to monitoring the activity of automated systems or controlling the vehicle.

- Poor take over performances, mode confusions, and inadequate trust were observed in partially automated vehicles.

- Interfaces indicating the reliability of automation have been shown to improve attention allocation, mental models and trust calibration.

- Multimodal interfaces have beneficial effect of attention allocation and potentially on mode awareness.

CHAPTER 2 – AUTOMATION MODE AWARENESS AND MULTIMODAL INTERFACES IN AUTOMOBILES:

a Systematic Review, Meta-Analysis, and a New Approach Based on the Signal Detection Theory

This theoretical chapter aims to provide a state of art of the effect of interface modality on mode awareness. It was the object of a review article submitted to the journal Cognition, Technology, and Work. It was reformatted for the purpose of this manuscript. The formatting consisted of reducing the introduction section to avoid redundancy with Chapter 1, and make consistent use of terms concerning driving automation.

Monsaingeon, N., Caroux, L., Langlois, S., & Lemercier, C. (submitted). Automation mode awareness and multimodal interfaces in automobiles: A systematic review and meta-analysis.

Résumé

L'objectif de cette étude est de fournir un état de l'art de l'effet de l'interface sur la conscience du mode et de proposer d'améliorer les mesures de la conscience du mode par la théorie de la détection des signaux. Pour assurer une conduite automatisée sécurisée, les conducteurs doivent être conscients du mode des systèmes automatisés. L'ajout d'autres modalités sensorielles aux interfaces visuelles classiques pourrait augmenter la conscience du mode grâce à la redondance d’informations. La théorie de la détection des signaux offre la possibilité de calculer des indices qui reflètent la capacité des participants à discriminer les différents états des systèmes automatisés. Une revue systématique de la littérature a été réalisée afin d'identifier la modalité sensorielle utilisée pour chaque interface, les techniques employées pour évaluer la conscience de mode et l'effet de la modalité de l'interface sur la conscience de mode. Pour quantifier l'effet des interfaces, les études précédentes ont été soumises à une méta-analyse. Afin d'évaluer la sensibilité des interfaces à provoquer des estimations correctes du mode en cours, des indices de détection signal ont été calculés. Les interfaces utilisant la vision centrale ou la vision périphérique, les signaux auditifs et les signaux vestibulaires peuvent améliorer la conscience du mode. Les interfaces en vision centrale semblent être plus efficaces que les interfaces multimodales. Grâce au calcul d'indices de détection du signal sur deux études, nous démontrons que certaines interfaces visuelles, considérées de sensibilité équivalente, étaient plus sensibles que d'autres. Les indices de détection des signaux semblent être une solution intéressante pour évaluer la conscience des modes. Les études futures devraient comparer les effets des interfaces auditives, vestibulaires ou haptiques et des interfaces en vision périphérique sur la conscience des modes.

Abstract

The purpose of this study is to provide a state of art of the effect of the interface on mode awareness and to propose to improve mode awareness measurements through Signal Detection Theory. To ensure safe automated driving, drivers have to be aware of the mode of automated systems. The addition of other sensory modalities to the classical visual interfaces could increase mode awareness due to redundancy gains. The signal detection theory offers the opportunity to calculate indices that reflect the ability of the participants to discriminate the different states of automated systems. A systematic review of the literature was carried to identify the sensory modality used for each interface, the techniques employed to evaluate mode awareness and the effect of interface modality on mode awareness. To quantify the effect of interfaces, previous studies were subjected to a meta-analysis. To assess the sensitivity of interfaces to cause correct estimations about active mode, signal detection indices were calculated. Interfaces using central vision or peripheral vision, auditory signals, and vestibular signals can improve mode awareness. Interfaces in central vision seem to be more efficient than multimodal interfaces. Thanks to the calculation of signal detection indices on two studies, we demonstrate that some visual interfaces, considered of equivalent sensitivity, were more sensitive than others. Signal detection indices appear to be a valuable solution to evaluate mode awareness. Future studies should compare the effects of auditory, vestibular or haptic interfaces and interfaces in peripheral vision on mode awareness.

1. Introduction

The present study focused on the effect of multimodal interfaces on drivers' mode awareness. Particular attention was paid to studies that had investigated the influence of interface modality on participants' ability to identify a system's active mode or state. We report the results of a systematic review of the literature on multimodal interfaces and mode identification. Mode awareness measurements employed in the literature are also reported and a lack of consideration for different possible incorrect identifications of modes is highlighted. Then, a meta-analysis was performed to confirm and quantify the findings of this review. Due to differences of measurement of mode awareness, a limited number of studies were included in the meta-analysis. Finally, we propose an approach based on Signal Detection Theory (SDT) to investigate mode awareness more thoroughly and to quantify it systematically. Signal detection indices were calculated to reflect the capacity of drivers to discriminate between the different modes of automated systems.

Mode confusion refers to a situation in which a user thinks a system is in a different mode than is actually the case (Baltzer et al. 2017). For example, a driver may estimate that the LCA is active when in fact it is not. When a mode confusion occurs, the user may commit a *mode error*, involving the execution of an intention that is adapted to one mode when in fact the system is operating in a different mode (Sarter et al. 1997). For example, the driver may not turn the steering wheel in a curve, owing to a mistaken estimation that the LCA is active. In this example, the driver did not detect when that LCA suspended, which can be considered as an *omission error* because the human operator failed to takeover when needed. This is the type of mode confusion that is usually reported in studies (Banks et al. 2018). However, another type of mode confusion must be considered (Janssen et al., 2019). In a situation in which the LCA stays active in a curve, the driver may turn the steering wheel, owing to a mistaken estimation that the LCA is inactive. This type of mode confusion can be considered as a *commission error*, or a false alarm, because the driver failed to detect that automated systems are in charge of driving. This can result in contrary actions performed by the human operator and the automated systems. In some vehicles, an active LCA becomes inactive after a certain strength is applied to the steering wheel, giving back the control of the trajectory to the driver. If it happens during a curve, it could cause the vehicle to swerve significantly.

Mode confusions and mode errors are related to a deficient mode awareness (Kurpiers et al. 2020). *Mode awareness* refers to both an understanding of how automated systems work, which relies on mental models (i.e., driver's representations of how these systems function), and an awareness of these systems' state of activation (Monk, 1986). Mode awareness is a subcategory of situational awareness, in that it shares the same features, namely perception, comprehension and projection (Endsley, 1995). Both components of mode awareness are influenced by the interface. Information provided by the interface can increase drivers' understanding of how the systems function, and therefore improve their mental models. It can also make them aware of each system's activation. For this to happen, drivers need to detect the signal corresponding to the current mode of automation. The detection of this signal by the driver will lead to the correct or incorrect estimations about the system's actual mode. The interface should aim to induce an estimation as close as possible to the actual state of automation in order to induce an adequate mode awareness. To evaluate the capacity of an interface to induce correct estimations about automation's state, we propose to exploit the *Signal Detection Theory* (SDT; Janssen et al. 2019).

SDT takes root in target detection of radars and was developed more extensively in psychology. Its objective is to evaluate the sensitivity of humans to discriminate between signals and noise (Stanislaw & Todorov, 1999). The number of correct and incorrect detections of a signal, depending on the presence and absence of the signal, are merged into indices. Such indices were used in diverse domains, such as acceptability judgments (Huang & Ferreira, 2020), vigilance in a monitoring task (Craig, 1979), but also in human factors and ergonomics (Caroux et al., 2018). Applied to the domain of automated driving, Janssen et al. (2019) suggested that it would allow the investigation of the identification of the state or mode of automated systems. Omission and commission errors would be exploited to characterise the estimation the drivers make on the state of automated systems. Through the calculation of indices, it would be possible to evaluate the sensitivity of drivers to distinguish the modes of automation, in relation to the information presented by an interface. Interfaces could easily be compared with a common measure that takes into account all possible estimations, right or wrong, about the state of automated systems.

The majority of today's vehicles use the visual modality to inform drivers about activation modes. However, humans have limited visual resources (Broadbent 2013; Wickens 2008), and as driving and supervising automated driving are mainly visual tasks, too much visual information from the interface can cause cognitive overload. According to Wickens's (2008) model, it is possible to distribute the demands of a task across the different attentional resource channels, namely focal vision, peripheral vision, hearing, and touch. Distributing information about activation modes across the different types of resources available to drivers would therefore make it possible reduce cognitive load. Following this logic, many multimodal interfaces targeting multiple attentional resources have been developed.

The purpose of the present study was to review research on the effect of multimodal interfaces on mode awareness, in order to determine whether multimodal interfaces are more efficient than visual-only interfaces when it comes to increasing mode awareness. To this end, we conducted a systematic review of the literature on mode awareness and the sensory modalities used in interfaces, listing the types of interfaces that have been studied, as well as the techniques that have been used to assess mode awareness. To quantify the effect of multimodal interfaces, we then undertook a meta-analysis comparing the effects of multimodal versus visual interfaces on mode awareness.

2. Systematic Review of the Literature

Mirnig et al. (2017) carried out a systematic review of the literature and patents concerning the different sensory modalities used to convey information in automated vehicles. One key finding was that different sensory modalities can be used to inform drivers about automation modes. Information can be conveyed in the visual modality (i.e., through texts, symbols and colors), but also in the auditory modality. However, these different modalities were not compared on efficiency. In a closely related study by Zhang et al. (2019) on highly automated vehicles (Levels 3 and 4), 129 studies were subjected to a meta-analysis that measured takeover time, which corresponds to the interval between the takeover request and the actual takeover. One of the main findings of this meta-analysis was that takeover time was shorter when a visual request was accompanied by an auditory or vibrotactile request, than when it was used on its own. This result highlights the potential usefulness of multimodal interfaces to convey alerts without disturbing the participant's visual attention. Highly automated vehicles included in Zhang et al. (2019) meta-analysis usually had two operating modes: manual driving or automated driving. However, mode errors often occur when there are numerous available modes, each with different operating states, as it is the case of partially automated vehicles. A systematic review of the efficiency of interfaces with different sensory modalities should include vehicles with at least partial automation, in order to design interfaces that ensure safe driving. The goal was thus to undertake an exhaustive review of the effect of interface on mode awareness. To this end, we formulated a research question and collected publications. These were sorted according to their degree of relevance to the research question. Information such as the type of interface and research results, was then extracted from the papers, compiled, and discussed.

2.1. Method

2.1.1. Framing the Research Question

The research question, formulated in accordance with Participant Intervention Comparator Outcomes (PICO; Schardt et al., 2007), was formulated as follows: Is a multimodal interface more efficient than a visual-only interface for promoting mode awareness among drivers?

2.1.2. Information Sources and Search Strategy

The systematic review followed PRISMA principles (Moher et al., 2009). Academic publications were gathered via Google Scholar, EBSCOhost, and Web of Science at the end of 2020. The references of these articles were then examined, to ensure that no publications had been missed. Keywords of related studies were gathered, and three groups of common keywords were identified. The first group concerned the types of vehicle and automation. The second group concerned the interface. The third group concerned mode identification. We used the Boolean operators AND and OR to connect the three groups of keywords. Within each group, keywords were connected with the operator OR. The three groups of keywords were connected with the operator AND. The following keyword search is an example of the syntax we used for Google Scholar: ("Automated driving" OR "conditional automation" OR "partial automation") AND ("human-machine interface" OR "driver-vehicle interface" OR "feedback") AND (“mode awareness” OR “mode errors” OR “mode confusion”). The research was conducted without any limitation of date. Publication dates ranged from 1988 to 2019. We gathered a total of 218 articles: 199 from Google Scholar, 8 from Web of Science, 6 from EBSCOhost, and 5 from references and personal sources.

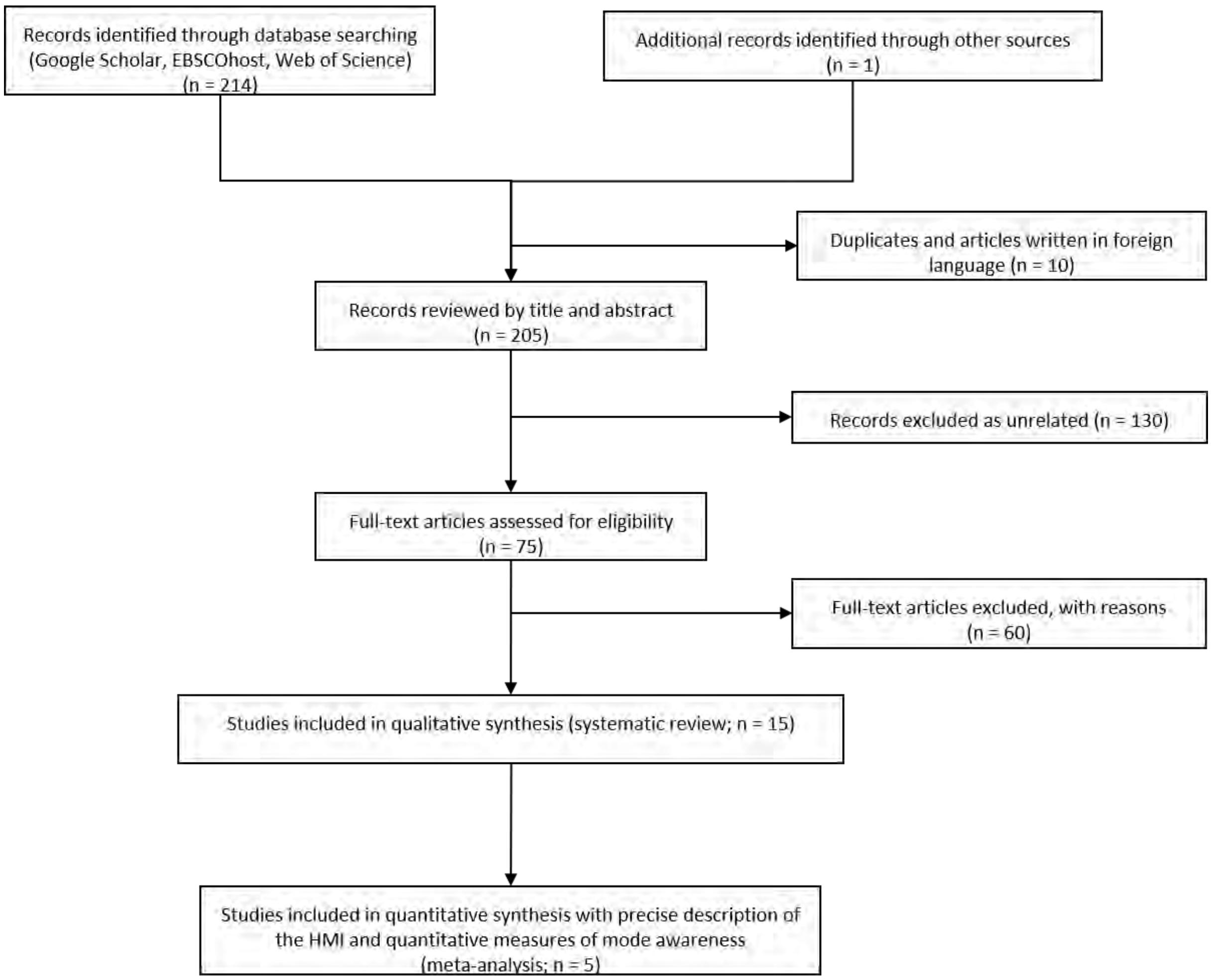
2.1.3. Exclusion Criteria

We applied the following exclusion criteria to the 218 articles: irrelevant for the study, does not have the relevant intervention (interface not mentioned), does not have a relevant comparison group, mode awareness is not assessed, written in a foreign language, and duplicate of another article. After reading the abstract, 140 records were removed because unrelated to the topic, duplicate of other studies, or written in foreign language. After reading the full texts, 60 studies were excluded with reasons, and 15 studies were included in the systematic review (see Figure 7 for PRISMA flowchart).

Figure 7

PRISMA flow diagram representing the systematic search strategy, including the identification, screening and inclusion of relevant studies.

Studies included in systematic review (n = 15); Studies included in meta-analysis (n = 5).



2.1.4. Data Extraction and Preparation

The following information was extracted from the 15 remaining papers: participants' mean age, level of vehicle automation, and whether the experiment took place on the road or in a simulator. We then extracted information about interface modalities. Table 2 describes the information collected from the studies.

Table 2

Variables extracted from the studies included in the systematic review.

Study variable	Coding	Description
Age	Years	Participants' mean age
Level of automation	0 = L1; 1 = L2; 2 = L3	Level of automation defined by SAE (2018)
Field of study	0 = Low fidelity simulator; 1 = High fidelity simulator; 2 = onroad study	Degree of realism. Lowest degree of realism = low fidelity simulator; highest degree of realism = real-life situation
Information displayed in focal vision, on the instrument's cluster	0 = No; 1 = Yes	Activated mode visually displayed on the cluster
Auditory modality for mode awareness	0 = No; 1 = Yes	Auditory signal of state of mode
Pitch motion	0 = No; 1 = Yes	Pitch motion from back to front to indicate mode of automated systems
Roll motion	0 = No; 1 = Yes	Roll motion from left to right or right to left to indicate mode of automated systems
Visual on the steering wheel	0 = No; 1 = Yes	Visual information on activation mode via LEDS on steering wheel
Peripheral vision	0 = No; 1 = Yes	Visual information on activated mode via LED stripes at the top of dashboard
HUD	0 = No; 1 = Yes	Head-up display indicating activated mode

2.2. Results

The systematic review yielded 15 papers. Coding of these papers revealed that different sensory modalities were used to inform the drivers of automated vehicles about activation modes (see Table 3). The vast majority of interfaces were visual only, with information about the activation mode presented for focal vision, on the instrument's cluster (n = 9). Drivers therefore had to glance at the cluster to receive this information. The second type of interface presented information about the activation mode in peripheral vision, through LEDs on the steering wheel (n = 1). The remaining studies used multimodal interfaces. Most multimodal interfaces combined the auditory modality with focal vision to inform drivers about mode transitions (n = 3). The last type of multimodal interface used focal vision and vestibular information via pitch or roll motion to inform drivers about activated modes (n = 2). Each type of interface is described in Table 4 below. All the studies took place in a simulator or on the road, and the participants' task was simply to drive.

Table 3

Table summarizing sensory modalities used by interfaces in each study to inform drivers about activation modes.

Authors	Focal vision	Peripheral vision	Auditory	Vestibular stimulation
Baltzer et al. (2017)	X			
Wang and Soffker (2019)	X			
Naujoks et al. (2017)	X			
Forster et al. (2016)	X			
Wandtner (2018)	X			
Lee and Ahn (2015)	X			
Eom and Lee (2015)	X			
Furukawa et al. (2003)	X			
Horiguchi et al. (2006)	X			
Belderbos (2015)	X	X		
Banks et al. (2018)	X		X	
Endsley (2017)	X		X	
Feldhütter et al. (2018)	X		X	
Cramer et al. (2018)	X			X
Cramer and Klohr (2019)	X			X

Table 4

Description of the studies integrated in the systematic review of literature.

Authors	Year	Levels of automation	Type of study	Driving environment	Sample size	Studied interface	Measure	Results
Baltzer et al.	2017	L0, L2	Assessment	Simulator	20	Visual interface on the instrument's cluster meant to be seen in focal vision, displaying alerts about events on the road.	Rating of mode awareness on Likert-like scale.	Better mode awareness with the visual interface.
Banks et al.	2018	L0, L1, L2	Observational	On road	12	Visual interface on the instrument's cluster meant to be seen in focal	Frequency of verbalized mode confusions.	Several verbalized mode confusions with the visual and auditory interface.

						vision, and auditory signals.		
Belderbos	2015	L0, L3	Comparative	Simulator	15	Visual interface with LEDS, meant to be perceived in peripheral vision	Questionnaire assessing mode confusions.	Interface with LEDs caused fewer mode confusions than an interface in focal vision.
Cramer et al.	2018	L2	Comparative	On road	36	Pitch motion interface.	Effect of pitch motion on awareness of activated mode rated on Likert-like scale.	High ratings of effect of pitch motion.
Cramer & Klohr	2019	L2	Comparative	Simulator	39	Roll motion interface, indicating vehicle's intentions.	Effect of roll motion on awareness of activated mode rated on a Likert-like scale	High ratings of effect of roll motion mode awareness.
Endsley	2017	L0, L1, L2	Observational	On road	1	Visual and auditory interface.	Number of mode confusions reported by the participants.	Author experienced mode confusions and mode errors

								with the visual and auditory interface.
Eom & Lee	2015	L0, L2	Comparative	Simulator	40	Visual interface displaying different states of modes.	Questions on current activation mode (number of correct answers) and subjective feelings of mode confusion.	More mode confusions were observed with the interfaces displaying 4 or 5 activation modes versus only 3.
Feldhuetter et al.	2018	L0, L2, L3	Comparative	Simulator	45	Visual and auditory interface.	Gaze, experimenter's rating of mode awareness on a Likert-like scale, interviews, participants' descriptions of how the system functioned	Manual driving phases inserted between phases of partially or highly automated driving did not affect mode awareness. Difficulty differentiating between modes was reported with the

								visual and auditory interfaces.
Forster et al.	2016	L3	Comparative	Simulator	6	Visual interface on the instrument's cluster displaying perception and vehicle's intentions.	Number of unnecessary deactivations of automation and questions probing comprehension of the system	The visual interface was mostly well understood. Drivers could not distinguish between the preparation and execution of the system's overtaking action.
Furukawa et al.	2003	L1	Comparative	Simulator (low fidelity)	40	Visual interface on the instrument's cluster displaying transitions between modes.	Percentage of participants who were unable to predict the system's behaviour	No differences in the number of mode confusions between the visual interface in which the states overlapped and the regular interface.
Horiguchi et al.	2006	L1	Observational	Simulator (low fidelity)	35	Visual interface on the	Mode confusion rate assessed by asking	Mode confusions more likely to occur

						instrument's cluster displaying mode states.	about the current activated mode.	when modes were closely related.
Lee & Ahn	2015	L0, L1	Comparative	Simulator	10	Visual interface on the instrument's cluster displaying mode states.	Mode confusion rate assessed by asking about the current activation mode.	The visual interface displaying more activation modes allowed drivers to make fewer mode confusions.
Naujoks et al.	2017	L4	Expert assessment	Simulator	6	Visual interface on the instrument's cluster displaying intentions of automated systems.	Questions probing comprehension of the system	Experts understood the color coding of the vehicle's intentions. Drivers could not distinguish between the preparation and execution of the system's overtaking action.

Wandtner	2018	L0, L3	Comparative	Simulator	36	Adaptive interface on the instrument's cluster displaying information when the driver is available.	Rating of system's limits on Likert-like scale.	All three interfaces led to good mode awareness, and did not differ between each other. The two adaptive interfaces led to a better awareness of the system's limits than the basic one.
Wang & Soffker	2018	L0, L1, L2, L3	Comparative	Simulator	38	Visual interface on the instrument's cluster meant to be seen in focal vision, displaying perceptions of the vehicle's events detection on the road.	Rating of mode awareness on Likert-like scale.	Collaboration between driver and vehicle enhanced by driver's awareness of upcoming events on the road and current activated mode.

2.2.1. Interface Using the Instrument's Cluster in Focal Vision

Displaying information in focal vision has many advantages over other sensory modalities. One major advantage is that it is possible to display semantic information. For example, it is possible to display information about both current and future activation modes. In Baltzer et al. (2017), a visual interface that provided information about events on the road and potential suspensions of automated systems was compared with a classic interface that did not. More specifically, the authors assessed the impact of these interfaces on mode awareness in a Level-2 vehicle via questionnaires. Results revealed that mode awareness was better when events on the road were flagged up. This created a better collaboration between the driver and the automation. For their part, Wang and Soffker (2019) tested an interface designed to enhance driver-vehicle interaction that featured visually displayed information on the instrument's cluster related to the road scene, such as vehicles detected around d the ego-vehicle. It also displayed clear information about the current activated mode and future mode transitions. After driving on a simulator with L1, L2 or L3 automated systems, participants answered questions on automation and situational awareness. Results revealed benefits of the interface on automation and situational awareness. The collaboration between the two actors (i.e., driver and vehicle) was enhanced by the fact that the driver was aware of upcoming events on the road and the activation mode. This allowed the drivers to anticipate possible mode changes.

Several studies looked at ways of displaying more complex visual information for drivers. If drivers are shown information about their vehicle’s intended actions, they should be able to anticipate potential mode transitions. Naujoks et al. (2017) designed a visual interface that could display the vehicle’s future actions, and it was tested in a simulator study with Level-3 automated systems (Forster et al., 2016). A color code was used to indicate what the system perceived and what it intended to do. Intentions to overtake other vehicles were displayed, as well as different degrees of imminence of transitions between modes. Comprehension of the system’s functioning was assessed with a questionnaire after a driving session featuring one of several scenarios. Results indicated that the interface was mostly well understood. Participants understood the color coding of the vehicle’s intentions. However, they were unable to distinguish between the preparation and execution of the system’s overtaking action. Therefore, the visual differences between the two events on the interface may not have been clear enough. Other studies took the driver’s ability and availability into account. Wandtner (2018) assessed the effect of visual adaptive interfaces that monitored the driver’s state when using Level-3 automated systems. Two adaptive interfaces were compared with a basic interface. In a first adaptive interface, the cluster changed when switching from one mode to another, and the reason for overtaking was displayed (e.g., roadworks). The second adaptive interface was similar to the first, but considered the driver's availability by assessing the latter's engagement in a secondary task. Mode awareness and awareness of the system's limits were assessed with questions after the driving sessions. Results revealed that all three interfaces led to efficient mode awareness, and did not differ from each other. However, the two adaptive interfaces led to a better awareness of the system's limits than the basic one. This suggests that clear distinctions between modes are needed for the driver to understand the system's functioning. It also suggests that the availability of the driver needs to be taken into account to allow the latter to have a clear understanding of the system's functioning and current activation mode.

The driver's ability to understand how the systems work and to be aware of the activation mode depends on how the different states of a modes are displayed. In a study by Lee and Ahn (2015) that focused on Level-1 automation, a classic visual interface displaying two activation modes, namely *waiting* (e.g., *ACC* in green) and *activated* (e.g., *ACC* in green and set speed in blue), was compared with a new interface that presented four states of the modes: *armed* (e.g., *ACC* in red and set speed in red), *cancelled* (e.g., *ACC* in red and set speed in blue), *override* (e.g., *ACC* in yellow and set speed in blue), and *activated* (e.g., *ACC* in green and set speed in blue). The number of mode confusions was evaluated by comparing the believed state of the system with the actual one, after pausing the experiment. Results revealed that displaying more activation modes allowed the drivers to have better mode awareness and make fewer mode confusions. However, displaying too many states of the system can be detrimental to mode awareness. Eom and Lee (2015) assessed the effect of the number of states displayed for a Level-1 vehicle. A visual interface displaying three states of the system was compared with one displaying four or five states. The number of mode confusions was measured after a pause in the experiment. Mode confusions were more numerous when four or five states of the system were displayed rather than three. These results suggest that although drivers need to be informed about state of automation, but it is important not to overload them with too much fine-grained information. More importantly, the states of automation that are to be displayed need to be carefully chosen, so that the help the driver accomplish the driving task.

A study focusing on Level-1 automation (Furukawa et al., 2003) featured Interfaces in which two states of an ACC (*low speed* and *high speed*) overlapped. This meant that even if the speed of the vehicle was supposed to cause a transition from low to high speed ACC, the clusters continued to display the previous state. These interfaces were compared with a regular one in which the states did not overlap. The number of mode confusions was counted by asking participants what they believed the state to be. Results revealed that there was no difference in the number of mode confusions between the interfaces in which the states overlapped and the regular interface. Driver convenience may therefore be enhanced by displaying information that is more easily comprehensible and does not compromise their safety. To make interfaces and automated systems clearer for drivers, the relationship between a system's behaviour and its activated mode needs to be studied. Horiguchi et al. (2006) did just this. Their hypothesis was that mode confusion can occur when a Level-1 automated system behaves in very similar ways in two different states (e.g., if the suspended state and the deactivated sate both result in inactivity of the system). The number of mode confusions was assessed by asking participants what they believed the system's state to be. The interface displayed all the different states of the system. Results revealed that mode confusions were more likely to occur when states of the system were closely related. Therefore, displays showing different states of automation must indicate clear differences in the system's behaviour.

To conclude on interfaces in focal vision, their greatest advantage is that they can display complex information, thereby enhancing awareness of the activated mode and informing the driver's mental model of the functioning of the system. However, drivers need to glance toward the cluster to access this information. Other visual interfaces can transmit information without requiring the driver to look away from the road.

2.2.2. Interface in Peripheral Vision

To keep the eyes of the driver on the road, it is possible to use interfaces in peripheral vision. Belderbos (2015) studied LEDs embedded in the steering wheel. Their colors changed according to the mode, and the time remaining before takeover was also displayed. The reason for the takeover was displayed in a Head-Up Display (HUD). This information was also shown on the cluster. The effect of information in peripheral vision on mode confusion was compared with that of a baseline interface in which information was displayed solely on the cluster. Mode confusion was measured with a questionnaire after driving sessions featuring Level-3 automated systems. Results revealed that overall, drivers experienced significantly fewer mode confusions when information was provided in peripheral vision during a takeover. This suggests that information displayed in peripheral vision can allow drivers to understand mode transitions and act accordingly within short periods of time. Multimodal interfaces can also efficiently convey information about mode transitions through auditory signals, for example.

2.2.3. Multimodal Interface

Auditory and Visual Interfaces

Auditory signals can be alarming and trigger rapid reactions in drivers. However, most auditory signals used to inform them about mode transitions are only informative, and are not intended to alert them. We asked whether drivers perceive these messages, and whether they understand them and act accordingly. Two on-road studies assessed interfaces in focal vision and auditory signals in a commercially available vehicle (Banks et al., 2018; Endsley, 2017). The vehicle in question, a Tesla Model S, emits auditory signals when the Level-2 automated systems are activated or suspended: two tones going from low to high frequency for activation, and two tones going from high to low frequency for suspension. In the study of Banks et al. (2018), participants drove the vehicle and their verbal commentaries were classified. Mode confusions were inferred from these commentaries. Results revealed that mode confusions occurred when drivers mistakenly believed that the automated systems were activated when they were not. The authors interpreted these results as a lack of system transparency. Information about the state of the automated systems was not correctly understood. Endsley (2017) encountered similar issues with the very same vehicle. After driving it for a while, she experienced mode confusion when she failed to activate the automated systems and mistakenly assumed that it was activated. She also made mode errors when she failed to notice that the system had initiated a mode transition. For their part, Banks et al. (2018) also attributed their results to a lack of transparency. The visual interface was not clear enough, and the auditory signals were not systematically perceived. After driving with assistance for a while, the author experienced an out-of-the-loop phenomenon, which is a disengagement from the task caused by either inactivity or excessive confidence in the system. To avoid this phenomenon, some studies have investigated the effect of takeover alerts by a multimodal interface. In a simulator study, Feldhütter et al. (2018) used visual and auditory interfaces to inform drivers about activated modes in a vehicle equipped with Level-2 and Level-3 automated systems. Icons were displayed on the cluster. An auditory signal, composed of a single note (*gong*), indicated when automated systems were available. LEDs on the dashboard pulsed when a mode changed. The effect of the interfaces was not evaluated. One group of participants passed through a manual driving (i.e., Level 0) phase when changing from one mode to another (e.g., Level 2 to Level 3), while another group directly changed modes. Mode awareness was assessed by measuring fixation time on the road, with questionnaires and interviews at the end of the experiment. Results indicated that inserting manual driving phases into the transitions between partial and highly automated driving did not affect mode awareness. However, during the interviews, participants reported difficulty distinguishing between the highly and partially automated modes in the interface because they were not sufficiently different. Once again, this highlights the importance of clearly indicating which mode is currently active. Multimodal interfaces could be useful for flagging up important differences between modes, especially if they can quickly re-engage drivers using their vestibular system.

Vestibular and Visual Interfaces

Some studies have assessed the impact of vestibular interfaces on mode awareness. In the study of Cramer et al. (2017) examined the effects of multimodal (visual and vestibular) interfaces on mode awareness in Level-2 vehicles. The activation mode was displayed on the cluster, and the vehicle's pitch motion (i.e., movements back and forth when the automated systems were activated) was used to convey information about the vehicle's intentions. Drivers rated the degree to which the pitch motion made them aware of the activation mode on a 5-point scale. Ratings were high ($M = 4$), but the efficiency of the pitch motion to inform drivers about the activation mode was not compared with that of a simple visual interface. A second study used roll motion as vestibular information when automated systems were activated (Cramer & Klohr, 2019). The activated mode was also indicated on the cluster. When Level-2 automated systems were activated, roll motions of the vehicle were used to indicate the system's intentions. The vehicle moved from left to right to indicate the intention to overtake another vehicle. Drivers rated the degree to which this motion made them aware of the activation mode on a 5-point scale. Ratings were high ($M = 4$).

Different modalities have been tested in multimodal Interfaces, although the majority used auditory and visual information, with auditory signals used to indicate mode transitions. However, mode confusions were reported with these types of interface. Vestibular interfaces were used to inform drivers about activated modes and the vehicle's intentions. Mode awareness was reported as good, but as no comparisons were made with other Interfaces, these results could not be corroborated.

2.2.4. Literature Review Conclusion

The systematic literature review revealed that several types of interface are used to give drivers information about the activated mode (see Table 3). Visual-only interfaces were the most common ones, but multimodal interfaces were also found. The most common types of multimodal interfaces combined visual and auditory information to tell drivers about mode transitions. Other multimodal interfaces combined visual and vestibular (e.g., pitch motion) information. It is worth noting that no studies included in this systematic review reported the use of haptic information through the steering wheel to inform drivers about activated mode. However, several studies assessed the effect of the interface on mode awareness: either mental models or awareness of the activated mode. The tools used to measure the two dimensions of mode awareness ranged from eye tracking (Feldhuetter et al., 2018) to freeze probe techniques (Lee & Ahn., 2015). In order to quantify and assess the overall effect of multimodal interfaces on mode awareness, we undertook a meta-analysis.

3. Meta-Analysis

The goal of the meta-analysis was to check whether the results of the systematic literature review were confirmed by quantitative data. In other words, to assess the extent to which the modality used for the interface affected mode awareness. We assessed the modality of the interface, whether it was visual, auditory, haptic, or vestibular. The measures extracted from the articles were necessarily quantitative, and referred to mode awareness, mode confusion, mode errors, and mode awareness of activated mode. The analysis therefore consisted in comparing effect sizes reflecting the difference between two interfaces (e.g., multimodal vs. visual-only), in order to determine which type of interface had the greatest effect on mode awareness.

3.1. Method

3.1.1. Analysis Method

A random-effect model was chosen to analyse the effect of the interface because studies included in this meta-analysis used similar but not identical methods of measurement of mode awareness. In this type of model, summary effect sizes represent the mean of distribution of effect sizes (Borenstein et al., 2011). The calculation of summary effect sizes is based on Cohen's d (i.e., standardized mean difference) and pooled standard deviations (i.e., weighted mean standard deviation). The degree of heterogeneity in the data was estimated using the restricted maximum-likelihood estimator (i.e., $\tau2$; Viechtbauer, 2005). A high $\tau2$ ($\tau2 < 1$) reflects a high degree of heterogeneity in the data. In addition to estimated $\tau2$, Cochran's Q test for heterogeneity (Cochran, 1954) and the $I2$ statistic (Higgins & Thompson, 2002) are reported. In cases where heterogeneity was detected (i.e., $\hat{\tau}2 > 0$τ^2 > 0, regardless of the results of the Q test), a prediction credible interval for the true outcomes is also provided (Riley et al., 2011).

3.1.2. Study Characteristics

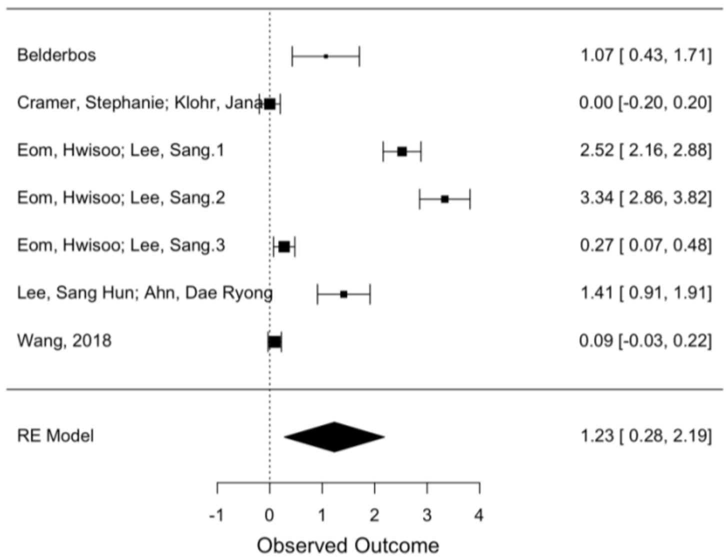
Five of the 15 studies in the systematic review were included in the meta-analysis, as these were ones that provided detailed descriptions of Interfaces, with pairwise comparisons and assessment of mode awareness or mode confusions with a quantitative measure. Three of these studies were published research articles, and two were chapters from an unpublished doctoral dissertation or master's thesis. It is debatable whether unpublished studies should be included in a meta-analysis. However, Rothstein and Bushman (2012) argued that it is possible, providing they meet the inclusion criteria, which in our case were the comparison of two interfaces and a quantitative measure of mode awareness. Publication date ranged from 2015 to 2019. Four of the five studies were conducted in a high-fidelity driving simulator with a motion platform, and one in a low-fidelity simulator. Regarding the assessment of mode awareness, one study used the freeze probe technique, one used ratings of correct mode identification, one used a Likert-like scale, and two used questions after the experiment.

3.2. Results

A total of five studies were included in the analysis. This number is small, but according to Davey et al. (2011), the mean number of studies included in meta-analyses is three. The observed standardized mean differences were all positive (range: 0.00-3.34). The estimated standardized mean difference based on the random effects model was $\hat{\mu} = 1.23$ (95% CI [0.28, 2.19]). The mean outcome differed significantly from zero ($z = 2.53$, $p = 0.011$). A forest plot showing the observed outcomes and the estimate based on the random effects model is shown in Figure 8. The left side of Figure 8 represents the reference interface, and the right side represents the experimental condition.

Figure 8

Forest plot showing the effect sizes and 95% confidence intervals (CIs) for comparisons between an experimental interface and a reference interface.
An effect size > 1 indicates a higher likelihood of better mode awareness with the experimental interface than with the reference one.

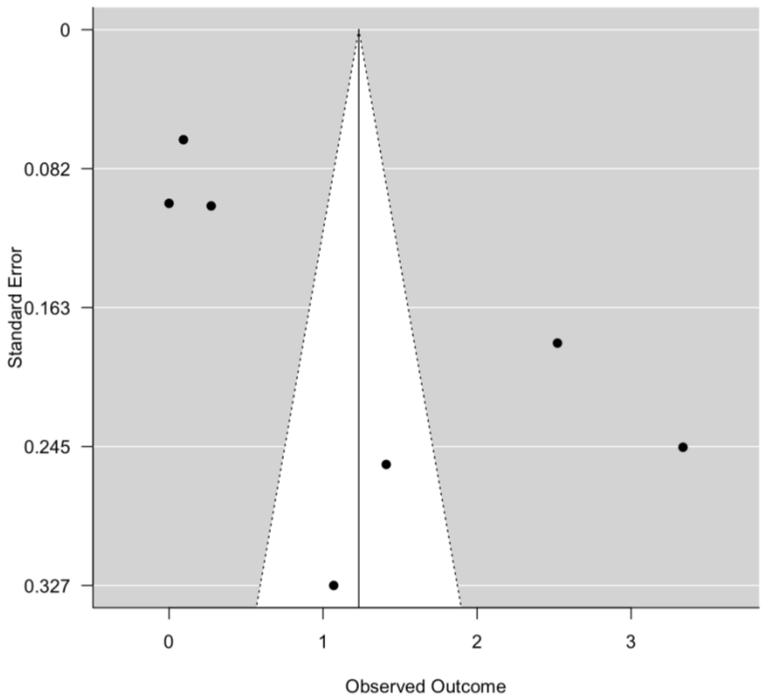


According to the Q test, true outcomes were heterogeneous, Q(6) = 335.184, $p < 0.001$, $\tau^2 = 1.62$, $\tau = 1.27$, $I^2 = 98.94\%$. The effects of the interfaces were not homogeneous, and varied according to the type of interface tested. The 95% prediction credible interval for the true outcomes ranged from 0.28 to 2.19. Hence, although the mean outcome was positive, in some studies the true outcome was close to zero. A funnel plot of the estimates is shown in Figure 9.

Neither the rank correlation nor the regression test indicated any funnel plot asymmetry (p = 0.136 and p = 0.100).

Figure 9

Funnel plot representing the homogeneity of the studies included in the meta-analysis.



3.3. Discussion

The purpose of this meta-analysis was to gauge the extent to which the modality used for the interface affects mode awareness. Studies of visual interfaces reported a greater effect on mode awareness than studies of multimodal interfaces. These results differed from those of the meta analysis of Zhang et al. (2019) concerning highly automated vehicles. These authors found that, compared with visual-only interfaces, multimodal interfaces combining visual with auditory or haptic stimulation allowed drivers to take over faster when a takeover was requested. However, in highly automated vehicles, the multimodal signals have different functions from those in partially automated vehicles. In highly automated vehicles, multimodal TORs call for immediate action and a reaction time is measured. In partially automated vehicles, information on active mode is permanently visually available on the instrument's cluster, while multimodal signals, such as auditory information, are used punctually to indicate mode transitions. The level of information processing is therefore different. Highly automated vehicles also differ from partially automated vehicles as they usually only have two modes (manual or automated driving). A more complex grammar of auditory or proprioceptive information is required and the drivers need to understand the meaning of each signal. Training is therefore necessary to be able to correctly associate the tones with the transition of the corresponding mode. This problem is less present with visual information because it can easily convey semantic information, which might be the reason why our results suggest that visual interfaces have more impact on mode awareness than multimodal interfaces. Another possible explanation is a lack of congruency between the visual and auditory or vestibular information, which failed to provoke a redundancy gain. In partially automated vehicles, drivers are continuously in a dual-task situation: driving or supervising while receiving information about mode transitions. The multimodal information therefore needs to be sufficiently attractive to capture their attention without distracting them from their driving or supervising task. It should also be clear enough to be understood by drivers with the shortest possible training. However, a very limited number of studies were included in this meta-analysis. One major reason for that is the wide variety of measurements of mode awareness, which often lack precision regarding the incorrect identifications of modes. A solution to take into account all possible correct or incorrect identifications of modes and to draw relevant measurements would rely on signal detection theory.

4. Mode Awareness Through The Lens of Signal Detection Theory

This section aims to describe an application of the SDT to evaluate the sensitivity of drivers to discriminate between the active and inactive states of automated systems. Through the calculation of indices, we aimed to determine the most efficient interface to indicate the state of automated systems, taking into account all possible estimations, right or wrong, about the state of automated systems. The study selection to apply the SDT is described below, as well as the indices calculation and the results of the analysis.

4.1. Method

4.1.1. Included Studies

The SDT was applied to two studies collected during the systematic review and which were included in the meta-analysis. Those studies were selected because they reported data that allowed to identify the number of omission errors (*Misses*), commission errors (*False Alarms*), correct identification of inactive states of automated systems (*Hit*) and correct identification of active states of automated systems (*Correct rejections*). The two included studies both used the Freeze Probe Technique and assessed the effect ACC states displays on mode confusions (Eom & Lee 2015; Lee & Ahn 2015).

4.1.2. Indices Calculation

Four type of responses were considered, depending on the estimation the driver had on the state of automation and the actual state of automation: Hit, Miss, False alarm (FA), and Correct Rejection (CR) (see Table 5 for an example adapted from Janssen et al. 2019). Based on the number of Hit, Miss, FA, and CR, the rate of Hit and FA were calculated (see Equation **1** and Equation **2**).

Table 5

Application of signal detection characterization based on ACC states and human agent estimation.

		Human agent estimations	
		Human in control	Automation in control
Non-human agent mode	ACC off	Hit	Miss
	ACC on	False alarm	Correct rejection

$$Hit\ rate = \frac{number\ of\ Hit}{number\ of\ Hit + number\ of\ Miss}$$

Equation 1: Hit rate calculation.

$$FA\ rate = \frac{number\ of\ FA}{number\ of\ FA + number\ of\ CR}$$

Equation 2: FA rate calculation.

Two main indices were calculated to investigate the discrimination between signal and noise: *d'* and β. In the context of automated driving, the *d'* index would represent the capacity of the participant to discriminate between the suspension and activation of automated systems (see Equation **3** for *d'*'s calculation). The *d'* usually varies from 0, representing random responses (i.e., Hit rate and FA rate both equal 0.5), to 4.65, representing a close to perfect discrimination between signal and noise (i.e., Hit rate equals 0.99 and FA rate equals 0.01). An important *d'*, close to 4.65, reveals that the driver is able to discriminate efficiently between the active and the inactive states of automation. The index β reveals the response bias, which is the tendency of the participant to respond more often that automated systems are active or inactive (see Equation **4** for β's calculation). A positive β reveals that the participants tend to be *conservative* and estimate that automation is more often active. A negative β reveals that the participants tend to be *liberal* and that they estimate that automation is more often inactive (McNicol, 1972).

$$d' = Normalized\ (Hit\ rate) - Normalized\ (FA\ rate)$$

Equation 3: d' calculation

$$\beta = Exponential\ \left(\frac{Normalized\ (FA\ rate)^2 - Normalized\ (Hit\ rate)^2}{2}\right)$$

Equation 4: β calculation

When the previously introduced indices *d'* and β could not be calculated because *Hit* rate or *FA* are equal to 0 or 1, non-parametrical indices were used (Stanislaw & Todorov 1999). The most commonly used non-parametrical signal detection index is *A'*. It was used as a substitute of *d'*, it can range from 0 to 1, and usually varies from .5 to 1. The closer *A'* is from 1, the more sensitive the interaction between the driver and the automated systems (see Stanislaw & Todorov, 1999 for more details; Equation **5**). To estimate response bias, non-parametrical index B''_D were calculated. As for β, B''_D varies from -1 to 1 (Equation **6**). A positive B''_D reveals that the participants tend to be conservative and a negative B''_D reveals that participants are liberal.

$$A' = \frac{1}{2} + \frac{(Hit\ rate - \ FA\ rate)(1 + Hit\ rate - FA\ rate)}{4Hit\ rate(1 - FA\ rate)}$$

Equation 5: Non-parametrical A' calculation when Hit rate ≥ FA rate

$$B''_D = \frac{Hit\ rate(1 - Hit\ rate) - FA\ rate(1 - FA\ rate)}{Hit\ rat(1 - Hit\ rate) + FA\ rate(1 - Fa\ rate)}$$

Equation 6: Non-parametrical B''$_D$ calculation when Hit rate ≥ to FA rate

4.2. Results

4.2.1. Eom & Lee (2015) Study on the Number of Represented States of ACC

In this study, the responses of participants were reported, as well as the actual state of ACC (refer to Section 2. Systematic review of the literature for a detailed description of the study). To carry out the SDT, the data of this study was reproduced. Given that our review is focused on the effect of the interface, we only considered the results when participants glanced at the instrument's cluster. Depending on the actual state of ACC and the response of the participant, the logic described in Table 5 was applied (see Table 6 for an example on the interface with 5 states of ACC).

Based on this classification, the Hit ratio and the FA ratio were calculated for each tested interface, as well as the *d'* and β indices. For the 5 states interface, the estimated *d'* was 3.31, which can be valued as an important sensibility of the participants to discriminate between the active and inactive states of the ACC. The β result was 0.37, signifying that participants tended to be conservative and to respond more often that the ACC was in inactive states. For the 4 states interface, the discrimination between active and inactive states of ACC was very good, with a *d'* of 4.26. With this interface, participants tended to be less conservative (β = 0.11) than with the 5 states interface. Regarding the 3 states interface, the hit ratio was very high (0.98) and no FAs were observed, which made it impossible to calculate *d'* and β. This suggest that participants were perfectly capable of discriminating between the active and inactive states of ACC (see Table 7 for a summary).

Table 6

Classification of participants' answers and actual states of ACC for the 5 states interface.

Interface with 5 states of ACC			Participants' answers					
			Human in control			Automation in control		Total
			Off	Armed	Canceled	Override	Active	
Actual states of the ACC	Inactive ACC	Off	84	0	0	0	0	84
		Armed	0	25	11	3	4	43
		Canceled	1	9	137	0	0	147
	Active ACC	Override	0	6	2	27	0	35
		Active	0	0	3	1	87	91
		Total	85	40	153	31	91	400

Note. The answers corresponding to Human in control and Inactive ACC were considered as Hits. The answers corresponding to Human in control and Active ACC were considered as FAs. The answers corresponding to Automation in control and Inactive ACC were considered as Misses. The answers corresponding to Automation in control and Active ACC were considered as CRs.

Table 7

Hit/miss ration, FA/CR ratio, d' and β depending on the interface

	Hit ratio	FA ratio	*d'*	β
5 states interface	0.97	0.09	3.31	0.37
4 states interface	1.00	0.05	4.26	0.11
3 states interface	0.98	0.00	-	-

4.2.2. Lee and Anh's (2015) Study on the ACC Visual Interface

A similar experiment methodology to Eom & Lee (2015) was used in this study (see Section 2.2.1. for a detailed description of the study). The two tested interfaces represented different states of ACC. Interface 1 represented only two states of ACC and interface 2 represented four states. The same data correspondence as mentioned above was carried out with the reported data. The Hit ratio was high for the interface 1 (0.74) and for interface 2 (1.00). However, for both interfaces, no FAs were observed based on our classification, which prevented the calculation of *d'* and β indices. Non-parametrical *A'* signal detection measure was applied. With interface 1, the results revealed a very important signal discrimination (*A'* = 0.93). Participants were even more sensitive to ACC state changes with interface 2 (*A'* = 1.00). With both interfaces, participants didn't encounter FA so B''_D couldn't be calculated.

4.3. Discussion

The SDT allowed to take into account the four possible cases of mode identification and to calculate indices reflecting the capacity of participants to discriminate between the states of automated systems. The SDT was applied to only two studies because of insufficient data reporting from the other collected studies. This highlights a lack of consideration for the different possibilities of estimations regarding actual states of automated systems. In Eom & Lee's (2015) study, the 3 states interface was arguably the interface that induced the better discrimination of states of automated systems, followed by the four states interface and finally the 5 states interfaces. These results are in line with those of the authors. However, authors didn't observe a significant difference between the four-state and five-state interfaces. Our results revealed that the four-states interface induced better sensitivity than the five-states interface, but that response bias was also less important and participants tended to estimate that automated systems were inactive more often. Regarding Lee and Ahn's (2015) study, our results are close to those of the authors. The non-parametrical evaluations of sensitivity show that interface 1 provides better discrimination than interface 2. However, the difference does not seem to be as large as that observed by the authors. The SDT approach allowed us to further investigate mode awareness by calculating the sensitivity of participants to discriminate the states of automated systems depending on the interface. The sensitivity of participants differed between the tested visual interfaces. A similar method could be applied to multimodal interfaces to evaluate their efficiency to induce adequate mode awareness.

5. General Discussion

5.1. Main Findings

The threefold aim of the present study was to review research on the effect of interface on mode awareness, to compare the effects of multimodal versus visual-only interfaces on mode awareness, and to apply the SDT to take into account all possible false estimations of modes. To this end, we conducted a systematic literature review, a meta-analysis, and a signal detection analysis. The systematic review allowed us to highlight several findings in the literature. First, the relationship between visual interfaces and mode awareness has been more extensively studied than that between multimodal interfaces and mode awareness. The main advantage of visual interfaces is that they convey complex information on the cluster, acquainting drivers about the active mode through icons and texts, and contributing to their mental models by providing contextual information. The results of the meta-analysis revealed that the difference between visual interfaces can be more important than the difference between multimodal interfaces. The design of visual interfaces seems to importantly impact mode awareness. The signals exploited in current multimodal interfaces might require training to be understood. The multimodal information should complete visual information in order to create a redundancy gain. Finally, the SDT analysis revealed that visual interfaces with four displayed states of automation induced a better discrimination of states of automation, but also a more important tendancy of participants to estimate that automated systems are inactive. This approach allowed to further explore the effect of visual interface on mode awareness. The effect of interfaces on the discrimination of states of automated system was quantifyied, providing a more accurate measure of mode awareness.

According to the multiple resources model's assumptions, there are several resource channels, allowing different types of information to be processed at the same time. Multimodal interfaces should enable the cognitive load caused by driving or supervising and monitoring the automated systems to be more widely distributed. In turn, this should allow drivers to improve their mode awareness. However, the results of the systematic literature review were mixed, as mode errors were observed for interfaces providing both visual and auditory feedback. Interfaces using central vision and peripheral vision seemed to improve mode awareness, as did interfaces using vestibular feedback. These results tend to prove that multimodal interfaces can efficiently distribute the cognitive load caused by driving or supervising and thus improve mode awareness. When we compared the effects of multimodal interfaces versus visual-only interfaces on mode awareness in our meta-analysis, the former appeared to be less efficient than the latter. Assuming that the information was not redundant, this result contradicts the multiple resources model, insofar as multimodal interfaces should be more efficient in communicating information. This emphasizes the need for sensory modalities to be used more efficiently. For example, in Banks et al. (2018), auditory signals were emitted during transitions between activation modes.

The effect sizes observed when comparing visual interfaces could also be observed when comparing multimodal interfaces. By studying whether it is better to provide information about three or four activation modes, as Lee and Ahn (2015) did for visual interfaces, it would be possible to find out whether it is best to use auditory signals to inform drivers about activation modes or mode transitions. Moreover, each sensory modality has specific features. Auditory signals are discrete, meaning that they are mostly used to inform drivers about mode transitions. By contrast, vestibular feedback can be continuous, meaning that it can be present the whole time a system is active. Accordingly, sensory modalities could each be used differently, in order to exploit them to the full.

In sum, these results suggest that visual interfaces, which must be present in vehicles equipped with automated systems, need to be intelligently designed before adding other types of sensory stimulation. As persistent information on the instrument's cluster, the current mode should emulate clear and unambiguous representations for the drivers to be effective. It is not the same for auditory and vestibular signals that do not have intrinsic meanings. The signals might not be understood by new users who might need training to correctly associate the signals with the mode. Finally, in order to benefit from redundancy gain of multimodal interfaces, information in the different channels should be complementary. The visual and auditory information should be designed together, as the auditory signal can be imagined as the transition from one visual information to another. For example, the visual information could blink at the same time as a repeated beep is emitted during a transition of mode.

The heterogeneity of the data once again highlights a lack of consensus on the best methods for measuring mode awareness. In the studies we reviewed, mode awareness was assessed using a wide variety of techniques. In many studies, it was measured by assessing mode confusions: the authors froze the experiment and asked participants which activation mode they estimated the automated systems to be in. Other studies evaluated mode confusions by counting mode confusions that were verbalized while driving. Comprehension of the system was mainly assessed by administering questionnaires after the experiment. Finally, some studies investigated mode awareness with the help of eye-tracking techniques, by measuring the duration of fixations on the road. In Level-2 vehicles, drivers are supposed to monitor the activity of the automated systems or to be actively engaged in the driving task. In both cases, they have to look at the road. In Level-3 vehicles, drivers can perform secondary tasks, and are therefore not required to look at the road. If the duration of fixations on the road is longer for Level 2 than for Level 3, it can be assumed that drivers have accurate mode awareness (Feldhütter et al. 2018). The wide variety of techniques used to measure mode awareness affected the number of studies included in the meta-analysis.

5.2. Limitations

A clear limitation of this study was the small number of publications on the impact of interface on mode awareness. Only five studies were included in the meta-analysis, making it impossible to compare groups of studies that had assessed the effects of interfaces in the same sensory modality. In order to compare different categories of studies, each category must be composed of at least four articles (Fu et al., 2011). As a result, our meta-analysis could not account for significant differences in effect sizes depending on interface modality. Although multimodal interfaces have been extensively studied and shown to be efficient in the case of emergency automated systems (Ho et al., 2017) or highly automated vehicles (Zhang et al. 2017), more studies need to be conducted with partially automated vehicles.

Mode awareness has been studied from different viewpoints and with various methods. The results of some techniques, based on verbalization, could not be included in the meta-analysis, as there were no group comparisons, and no means were calculated. Some techniques involve the use of questionnaires, and others eye-tracking measures. There appears to be a need for consensus and for robust techniques that allow both the knowledge and awareness aspects of mode awareness to be assessed. Kurpiers et al. (2020) proposed a method that enables both the driver's mental models and awareness of the activation mode to be assessed using eye-tracking measures. A SDT approach completes this proposition, by allowing an in-depth investigation of the effect of interfaces on estimations of drivers regarding the mode of automated systems. A widespread application of SDT on several interfaces would allow to have a unified indicator of mode awareness and compare interfaces on a larger scale.

5.3. Conclusion and Recommendations for Future Research

The main conclusion of the present study is that multiple sensory modalities can be used in interfaces to inform drivers about the activation mode and mode transitions in partially automated vehicles. Interfaces in central vision have been extensively studied. Mode awareness can be improved by displaying the vehicle's intentions and future transitions between modes, as well as displaying the right number of activation modes. Peripheral vision can also improve mode awareness while keeping the road in central vision. Auditory signals coupled with visual information allows drivers to be informed about mode transitions, but have been associated with mode errors. Vestibular feedback allows drivers to be continuously informed about the current activation mode through pitch or roll motion, and therefore improves mode awareness. However, properly designed visual interfaces tend to have a greater impact on mode awareness than interfaces using other sensory modalities. Studies need to be conducted to determine how other sensory modalities could be used, be it to inform drivers about transitions between modes or to keep them continuously informed about active modes.

Visual information, displayed in central or peripheral vision, appears to be more suitable for keeping drivers informed about activation modes, whereas auditory information appears to be more suitable for informing them about mode transitions. Visual information in peripheral vision has the advantage of informing drivers whilst enabling them to keep the road in central vision. However, this raises the question of whether it is more efficient to present the two types of information in one or two modalities. Finally, techniques for measuring mode awareness need to be improved, in order to probe both knowledge about the functioning of the system and awareness of the activation mode. The proposition of Kurpiers et al. (2020) would allow for a comprehensive measure of mode awareness. SDT completes this proposition by allowing to unify measurement of mode awareness. To go further, the hidden Markov framework could be used to obtain a real-time measurement of mode awareness (Janssen et al. 2019). According to this method, the driver's behaviour could be used to predict mode awareness.

Points clés

- La conscience des modes peut être évaluée à l'aide de plusieurs techniques, telles que les questionnaires, la technique « Freeze-Prob », les mesures de fixation oculaire et les mesures comportementales.

- Les modalités utilisées par les interfaces pour informer sur le statut des systèmes automatisés étaient principalement visuelles en vision focale

- Certaines études ont testé des interfaces comportant des informations visuelles en vision périphérique, auditives ou vestibulaires. Aucune étude ne s'est intéressée au retour haptique pour améliorer la conscience du mode.

- Les tailles d'effet des comparaisons entre les interfaces visuelles étaient plus importantes que les tailles d'effet des comparaisons entre interfaces visuelles et multimodales.

- Les interfaces visuelles semblent être plus efficaces pour transmettre des informations complexes telles que les intentions d'automatisation, tandis que les interfaces auditives et vestibulaires semblent être plus adaptées pour informer sur les transitions de mode.

- La théorie de la détection du signal permet de considérer l’ensemble des confusions de mode possible et de calculer des indices de détection des interfaces

Key points

- Mode awareness can be assessed with multiple technics, such as questionnaires, Freeze-Prob technic, eye-tracking measures and behavioural measures.
- Modalities used by interfaces to inform on state of automation were mainly visual in focal vision
- Some studies tested interfaces with visual information in peripheral vision, auditory, or vestibular information. No studies investigated haptic feedback to improve mode awareness.
- Effect sizes of comparisons between interfaces with varying visual displays were more important than effect sizes of comparisons between visual with multimodal interfaces.
- Visual interfaces appear to be more efficient to transmit complex information such as intentions of automation, whereas auditory and vestibular interfaces seem to be more adapted to inform on mode transitions.
- The signal detection theory allows to consider all possible types of mode confusions and to calculate interface detection indices

Chapter 3 – The problematic of the present work

1. Summary of Literature Review

The previously cited literature points out that partially automated driving systems raises certain challenges. The drivers' attention needs to be allocated correctly between monitoring and controlling the vehicle. They need to have accurate mode awareness to know the limits of automation and to be aware of the state of automated systems. Finally, the trust they place in automated systems should be calibrated adequately to the system's capacity. Interfaces constitute the bond that unites the driver and the automated systems. Through the interface circulate information that can address these challenges. The purpose of this work is to design and evaluate interfaces that allow to allocate attention correctly, stimulate mode awareness, and induce adequate trust calibration. Multimodal interfaces indicating the limits of automation with peripheral vision appear to be relevant for that purpose.

Some commercially available partially automated vehicles are already equipped with multimodal interfaces. Studies were carried out with these vehicles and used interfaces composed of auditory signals in the form of earcons and visual information relative to the drivers' needs, presented in focal vision. Theses interfaces were associated to defective trust calibration and mode errors. Reliability information regarding automated systems is lacking in these vehicles. Yet, reliability information enables to orient the drivers' attention toward relevant elements when automated systems risk to suspend, increasing knowledge and awareness of state of automation, and helping to calibrate trust.

Considerations regarding the multiple resources model revealed that multimodal interfaces can complete reliability interfaces by informing on the state of automation without disturbing drivers. The systematic review of literature and meta-analysis presented in Chapter 2 reports that auditory interfaces indicating automation's modes were related to mode confusion involving. Regarding haptic interfaces, Chapter 2 highlighted that no study was carried to evaluate their effects on mode awareness. Finally, an important component of mode awareness and attention allocation is the drivers' representation of the system. Repeated exposure to automated systems and interfaces is required to build accurate mental models.

2. Gap in the Literature

Questions are still open regarding the effect of interfaces of existing partially automated vehicles on mode awareness and attention allocation. No comparison between multimodal driver-center interfaces and classical vehicle-center interfaces using only focal vision were done before. Such comparison would allow to understand the positive and negative effects of existing multimodal interfaces on attention allocation and mode awareness. On the basis of the results of a study doing such comparison, interfaces that overcome the shortcomings of the existing interfaces can be proposed.

Reliability interfaces discussed in Chapter 1 were mainly studied in degraded situations, where the drivers are somehow forced to comply to the reliability display. In real life situations, many environmental conditions can vary. It has not been yet investigated how drivers would comply to reliability information while environmental conditions vary from adequate to degraded. Would drivers blindly comply to reliability information without considering the quality of environmental condition or would they consider it in accordance with the environment? If drivers do comply to reliability information while considering the quality of environmental conditions, how should be designed the reliability interface? Several solutions exist but the more adapted in terms of attentional demand needs to be found.

The cause of mode errors related to auditory interfaces is not clear. Is it because of the usage of the auditory channel in addition to the visual one, or is it the design of the auditory signals that was not adapted? The sound design of these interfaces must be developed in such a way that the signals are effectively understood as representing automation modes. Secondly, a methodical evaluation of these signals before integration into the vehicle is necessary. This should ensure that the auditory interfaces induce an accurate mode awareness.

Regarding haptic interfaces, Chapter 2 highlighted a lack of research into their effects on mode awareness. Yet, the steering wheel is an obvious medium binding the driver and the automated lateral control. The ways of transmitting information through the steering wheel should be studied to identify the most appropriate ones to indicate the state of the automated systems.

Studies reported in Chapter 1 and Chapter 2 often report using multimodal interface without describing the development of each interface composing it. If the multimodal interface reveals not to be efficient, it is difficult to determine which elements of interface is to be blamed. The effectiveness of each element of a multimodal interface should be determined separately before integration to the multimodal interface. Finally, Chapter 1 highlights the importance of repeated exposure on the formation of mental models. The effect of interfaces on attention allocation and mode awareness is usually assessed after short periods of driving. In order to study the effect of a multimodal interface in an ecological framework, simulating the initiation of drivers after the purchase of a vehicle, a longitudinal study is needed.

3. Development of Interface

The literature review highlights either a lack of reliability information in interfaces or problems with existing multimodal interfaces. Therefore, controlled experiments studying each type of interfaces are needed to identify to what extent they influence attention distribution and mode awareness. For that purpose, several interfaces were designed in the context of CMI Project to compensate for the flaws of existing interfaces and gaps in the literature: an indicator of proximity to the limits of automation, a haptic interface in the steering wheel and an auditory interface. CMI's project and Renault's constraints were to use the instrument's cluster as a support for the purpose of indicating the proximity to the limits of automation, as it is a display that is present in all vehicles. It would allow to continuously transmit information on approach to the limits of automation in peripheral vision and to indicate the limits that are reached and the correct action to perform.

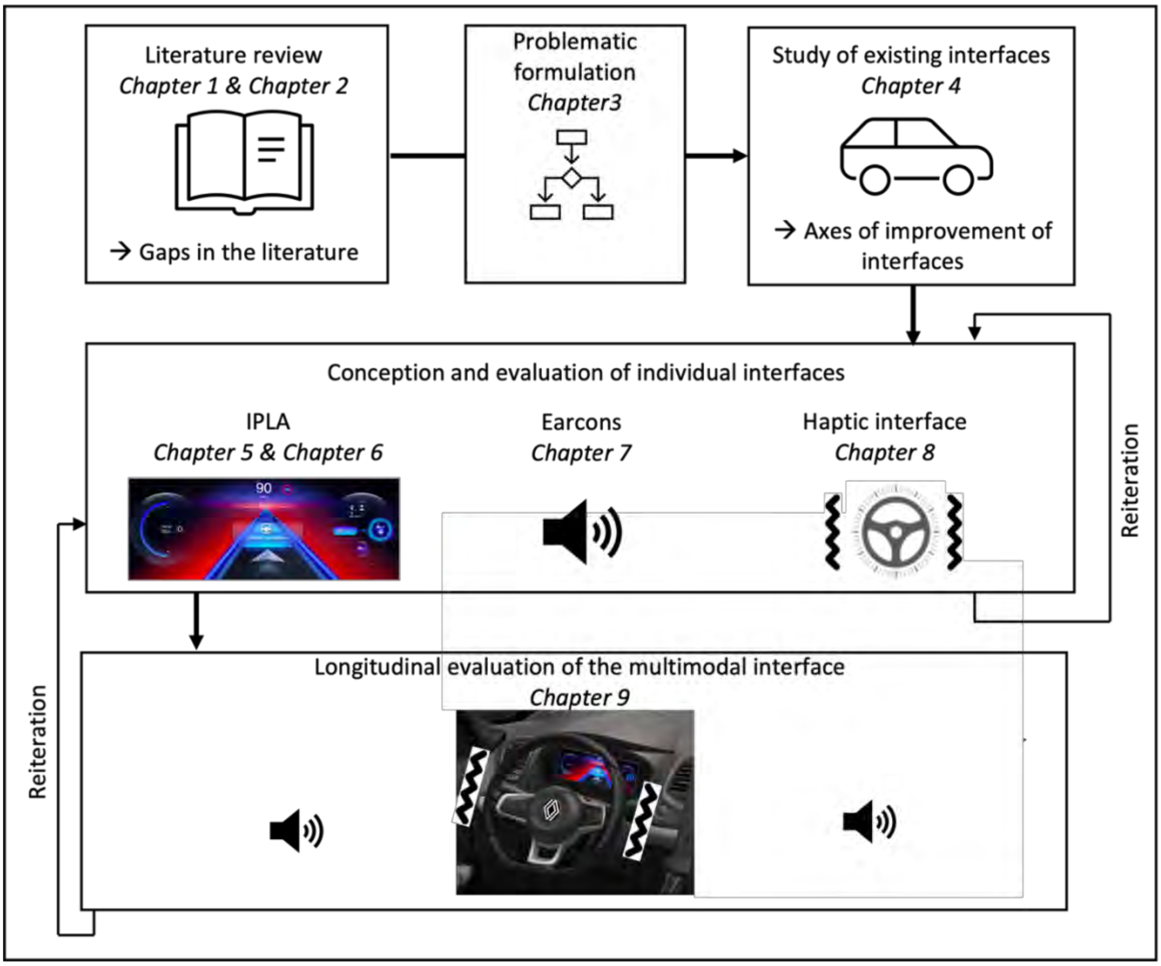
The haptic and auditory interface were developed in collaboration with collaborators of CMI Project using an iterative approach. The auditory interface was designed to be coherent with the sound design of CMI's Project. Two iterations of sound design were tested by members of the project before larger scale evaluations of their efficiency. The haptic interface was specifically designed to signal transitions of modes of LCA. Its design builds on earlier exploratory studies in Renault, which highlighted the potential of jerks and stiffness to inform on the mode of automation.

4. General Methodology

The problematic of this thesis was addressed following a methodology that aimed to study the effect of interfaces on mode awareness, attention allocation and trust in automation. It first consisted in studying the existing interfaces of partially automated vehicles. Then, the defaults of existing interfaces were highlighted, and specific elements of interfaces were designed to address these defaults. All elements of the interface were designed and evaluated to ensure that they fulfilled their purpose. They were then gathered in a multimodal interface to compare their conjugated effect with a classical visual interface using only focal vision (see Figure 10 for a description of the methodology).

Figure 10

Description of the methodology that was followed in this thesis to design and evaluate the efficiency of interfaces.



5. Plan of the Thesis

The capacity of multimodal interfaces to efficiently orient attention toward relevant elements of the interface and of the environment has been shown in Level-3 vehicles. Yet, the particularity of Level-2 vehicles is to monitor and understand the activity of automation. In the first study of the Experimental Section, an on-road study was carried out in which two interfaces were compared: a multimodal interface (i.e., presenting sounds and visuals in focal vision) center on the drivers (i.e., taking into account the humans' cognitive limitations), and a visual only interface center on the vehicle (i.e., displaying information about the vehicle's characteristics). The orientation of attention and knowledge about the system's characteristic should be better in a vehicle where the information depicts what the drivers need, which is the case in some existing vehicles. The results of this study, presented in Chapter 4, indicate that the multimodal interface center on the drivers led to better understanding of the vehicle's functioning but also more important visual demands for.

The first element of interface that was designed and evaluated was the indicator of limits of automation. The compliance of drivers to such indicator is to be clarified when several detrimental environmental conditions and prior knowledge vary. The presentation of proximity to the limits of automation and detrimental environmental conditions should impact the drivers' usage of automation. A study testing this hypothesis is presented in Chapter 5. Positive results of the presentation of an IPLA brought questions about which form should this indicator have. According to the literature, an IPLA presented in peripheral vision should induce accurate behaviours while not disturbing the drivers. The design of an IPLA aimed to be perceived in peripheral vision is presented and tested in Chapter 6. The results indicated positive effect of the IPLA on the anticipation of suspensions of automated systems but less appropriated actions. Based on these results, the design of the IPLA was improved and included in a multimodal interface.

The results of the systematic review of literature indicated that multimodal interface could have beneficial effects on mode awareness. Still, the quantitative data were not sufficient to prove their efficiency. Particularly, auditory interfaces have proven to be efficient in Level 3 vehicles but led to mode confusions in Level-2 vehicles. A method to ensure that the design of auditory interfaces induce correct mode awareness according to Situational Awareness Theory (Endsley, 1995) is presented in Chapter 7. It was applied on earcons that represented the hierarchy of levels of automation thanks to pitch, number of notes and timber, in order to evaluate their impact mode awareness. Based on the results presented in Chapter 7, the auditory interface was integrated to the final multimodal interface. Another interface element studied was a haptic interface. The absence of study investigating its effect on mode awareness in the systematic review highlights a lack of experimental data regarding the effect of these interface. A haptic interface was designed and evaluated in a simulator study to evaluate its effect on mode awareness. Positive results on quality and rapidity of detection of automation's suspension are presented in Chapter 8 and led this interface to be integrated in the multimodal interface.

The elements of interface tested in Chapter 5, Chapter 6, Chapter 7, and Chapter 8, which proved to be beneficial for attention allocation and mode awareness, were gathered in a multimodal interface indicating limits of automation with peripheral vision. This interface should induce better attention allocation, mode awareness, and calibrated trust in the automation than a visual only interface. This hypothesis was tested in a longitudinal simulator study and presented in Chapter 9. The results revealed that mental models regarding automation's limits and attention allocation were improved with the multimodal interface indicating limits. Driving performances after suspensions of automation were also improved to a more limited degree. More importantly, these improvements in attention, mental models, and driving behaviour were context dependent. Knowledge about automation's limits improved in some situation but not the quality of behaviours, and vice versa. Yet, trust in automation increased with the multimodal interface. This highlights the impact of the nature of information that is used by drivers depending on the context. This matter, along with the combined results of all chapters, is discussed in the General Discussion.

EXPERIMENTAL SECTION

The literature review presented in Chapter 1 and Chapter 2 led to the conclusion that partially automated vehicles raise challenges that interfaces could address. Multimodal interfaces can transmit information on state of automation without disturbing the drivers' central vision. Indicators of proximity to the limits of automation can improve mode awareness and mental models regarding automated systems' functioning. The following section will describe 6 studies investigating the effect of multimodal interface and reliability information of attention allocation, mode awareness and trust calibration. Some existing vehicles' interfaces might already have elements of response and will be studied first. Deficiencies and assets will be identified in exiting interfaces to develop the axes of work on which new elements of interfaces investigate. These new elements of interfaces will be designed and evaluated in separate studies. After that, all the new interface elements will be gathered into one multimodal interface and evaluated.

CHAPTER 4 – IMPACT OF INTERFACE DESIGN ON DRIVERS' BEHAVIOUR IN PARTIALLY AUTOMATED CARS: AN ON-ROAD STUDY

In this chapter, we investigate the effect of a multimodal interface on the understanding of automated vehicles' functioning. In the study presented in this chapter, two partially automated vehicles, available on the marketplace, were confronted. They differed in the orientation of their interface design. One interface was multimodal and represented information useful for the driver, while the other one was only visual and depicted information related to the vehicle's state. The vehicles were driven on-road for a short travel, considered as a first time use situation. The comprehension of drivers regarding the vehicle's functioning was assessed, which allowed to highlight benefits and costs related to the multimodal driver-centered interface. The conclusions of this study support the necessity to evaluate interfaces separately and to investigate them in longitudinal studies. The experimental study presented in this chapter was the subject of a research article published in the journal Transportation research part F: traffic psychology and behaviour. The article was reformatted for the purpose of this manuscript. The formatting consisted in making consisted usage of terminology and reducing introduction elements to avoid repetitions with previous chapters.

Monsaingeon, N., Caroux, L., Mouginé, A., Langlois, S., & Lemercier, C. (2021). Impact of interface design on drivers' behaviour in partially automated cars: An on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *81*, 508-521.

Résumé

Dans les véhicules partiellement automatisés, le conducteur et le système automatisé partagent le contrôle du véhicule. Par conséquent, le conducteur peut être amené à passer de l'activité de conduite conduite à la surveillance des systèmes automatisés. Cela peut avoir un impact critique sur la conscience de la situation du conducteur. L'interface est responsable de la collaboration efficace entre le conducteur et le système. Elle doit tenir le conducteur informé de l'état et des capacités des systèmes automatisés, afin qu'il sache qui est en charge de la conduite. La présente étude a été conçue pour comparer la capacité de deux interfaces avec des affichages d'informations différents pour informer le conducteur sur l'état et les capacités du système : une interface centrée sur le conducteur qui exploite une interface multimodale et une représentation exocentrique de la scène routière, avec et une interface traditionnelle centrée sur le véhicule qui exploite une interface visuelle. L'impact de ces interfaces sur les conducteurs a été comparé pendant une étude sur route. Les mouvements oculaires des conducteurs et leurs temps de réponse à des questions posées pendant la conduite ont été mesurés. Leurs verbalisations pendant la conduite ont également été transcrites et codées. Les résultats ont révélé des temps de réponse plus courts aux questions sur la vitesse du véhicule lorsque l'interface exocentrique et multimodale était utilisée. La durée et le nombre de fixations sur le compteur de vitesse étaient également plus élevés avec l'IHM centrée sur le conducteur. L'interface exocentrique et multimodale a aidé les conducteurs à comprendre le fonctionnement du système, mais elle était plus distrayante visuellement que l'interface traditionnelle. Les deux interfaces ont provoqué des confusions de mode. L'utilisation d'une interface multimodale peut être bénéfique et devrait être privilégiée par les concepteurs. L'utilisation d'interfaces auditives pour indiquer le niveau d'automatisation doit être explorée dans des études longitudinales.

Abstract

In partially automated vehicles, the driver and the automated system share control of the vehicle. Consequently, the driver may have to switch between driving and monitoring activities. This can critically impact the driver's situational awareness. The interface is responsible for efficient collaboration between driver and system. It must keep the driver informed about the state and capabilities of the automated system, so that he or she knows who or what is in charge of the driving. The present study was designed to compare the ability of two interfaces with different information displays to inform the driver about the system's state and capabilities: a driving-centered interface that displayed information in a multimodal way, with an exocentric representation of the road scene, and a vehicle-centered interface that displayed information in a more traditional visual way. The impact of these interfaces on drivers was compared in an on-road study. Drivers' eye movements and response times for questions asked while driving were measured. Their verbalizations during the test were also transcribed and coded. Results revealed shorter response times for questions on speed with the exocentric and multimodal interface. The duration and number of fixations on the speedometer were also greater with the driving-centered interface. The exocentric and multimodal interface helped drivers understand the functioning of the system, but was more visually distracting than the traditional interface. Both interfaces caused mode confusions. The use of a multimodal interface can be beneficial and should be prioritized by designers. The use of auditory feedback to provide information about the level of automation needs to be explored in longitudinal studies.

1. Introduction

The present study investigated how the design of the interface in existing partially automated vehicles influences drivers' understanding of the level of automation and the functioning of the system. In the present exploratory study, two vehicles were driven on road to evaluate the impact of two different interface in an ecological setting. Existing partially automated vehicles appear to have difficulty communicating the correct information to the driver. Drivers have been found to be less inclined to look at the road when using the Tesla Model S with Level 2 systems activated than they are when using a Level 1 or Level 0 (Gaspar & Carney, 2019). These authors found that drivers potentially over-relied on the system and were not fully aware of the role they still had to play. In a naturalistic study featuring the Tesla Model S, Banks et al. (2018) observed confusion between the levels of automation. Participants were filmed and recorded as they drove on the road for 40 minutes. The recorded behaviours were then analysed. At some point, several drivers let go of the steering wheel, in the mistaken belief that they had activated the Level 2 automation. They only realized that they were confusing modes of automation when they consulted the interface display. Mode confusions can lead to very hazardous situations, if the driver expects the vehicle to behave in a specific way and it does not (Sarter & Woods, 1995). For example, if drivers think that the vehicle is in charge of lateral control, they will not move the steering wheel in a bend, meaning that the vehicle ploughs straight into the side of the road. Banks et al. (2018) suggested that one possible reason for these mode confusions is that drivers place too much faith in the system, but hypothesized that the interface's lack of transparency is the most likely cause.

1.1. Mental Model

As the relation between the driver and the system changes according to the degree of automatization, so the interface has to change too. Carsten and Martens (2019) recently established goals that need to be achieved when designing interfaces for automated vehicles. The first goal is to ensure that the driver understands the capabilities of the vehicle and the level of automation that has been activated. To achieve this goal, interfaces need to adapt to the new constraints that automation places on the driving activity. In the present study, we compared two interfaces to evaluate which one came closer to meeting the first goal of Carsten and Martens (2019). Interfaces need to foster accurate mental models and avoid mode confusions. *Mental models* are the representations that humans have of a system's purpose, form, functioning, state, and structure (see Seppelt & Victor, 2020, for a more detailed definition). In the context of automated driving, drivers with an accurate mental model know when the automated system has been activated or suspended. They also understand which situations are appropriate for using the system. The information displayed by the interface needs to clearly be understood, so that the driver forms an accurate mental model at the very beginning of the interaction. To be understood, the interface must be tailored to the human's cognitive limits and to the specificity of driving an automated vehicle. Several cognitive models have been built to inform design in transportation and are discussed in the next section.

1.2. Design Solution to Improve Interaction with the Automated System

When driving, humans have to process a large amount of visual information from both the cockpit and the external environment. One model that is relevant to interface design because it takes account of humans' cognitive limitations is the multiple resource model developed by Wickens (2008). Currently used in the design of interfaces for complex environments that require simultaneous information processing, this model can predict the allocation of attentional resources to tasks performed simultaneously according to their qualitative characteristics. For the same amount of information, tasks in different sensory modalities (e.g., visual and auditory) are performed better than tasks in a single sensory modality. Accordingly, interfaces that use multiple sensory modalities to communicate should elicit a more effective distribution of attentional resources. The meta-analysis of Zhang et al. (2019) showed that interfaces providing a combination of visual and auditory or tactile information allow takeover time to be reduced in Level 3 vehicles. Interfaces using two or more sensory channels to inform the driver on the state of automation help the latter understand which actions have to be performed. Use of a similar interface in Level 2 vehicles could also improve interaction between the automated system and the driver.

1.3. Research Question and Hypotheses

The number of studies carried out on open roads with Level 2 vehicles is increasing (Banks et al., 2018; Endsley, 2017; Solís-Marcos & Kircher, 2019). So far, studies have investigated the impact on the driver of driving an automated vehicle at a subjective or cognitive level. However, none of them have compared different interface designs, even though some interfaces may be safer than others. According to the multiple resource model and previous studies of Level 3 vehicles (Wickens, 2008; Zhang et al., 2019), a driving-centered interface that takes the limitations of the human cognitive system into account by providing information via multiple resource channels (i.e., visual and auditory) should be more easily understood and elicit more accurate mental models, compared with a classic visual-only vehicle-centered interface. We therefore asked whether a multimodal interface centered on the driving activity and the limitations of the human cognitive system is more efficient than a visual interface centered on the state of the automation. By comparing the on-road usage of two existing vehicle interfaces, we sought to identify which interface characteristics help drivers to understand the capabilities and automation's state of their vehicle. We hypothesized that compared with the classic visual vehicle-centered interface, an exocentric and multimodal driving-centered interface allows the driver to have a better understanding of how the vehicle operates when the automated system is activated. This translates as better retrieval of information about the system, fewer visual fixations on the cluster, and fewer mode confusions. To test this hypothesis, participants drove one of two vehicles equipped with interface that fitted our description (i.e., driving-centered interface vs. vehicle-centered interface) on an open road while using automated systems.

2. Material & Method

2.1. Participants

We recruited 20 volunteers (19 men) aged 27-59 years (M = 40.80, SD = 8.53). They had held a driving license for a mean period of 21 years. The volunteers had no visual, even corrected (myopia, astigmatism, presbyopia), or auditory impairment. They were recruited among employees at Renault Group's Technocentre site in France. They were not paid, and they all signed an informed consent form. These volunteers were randomly assigned to one of the two vehicles (10 participants for each vehicle). The main criterion for recruitment was to have never driven the vehicle to which they were assigned. Seven participants had already driven a Level 1 vehicle before the experiment: two in the vehicle-centered interface group, and five in the driving-centered interface group. None of the participants had driven Level 2 vehicles before. All participants were familiar with an automatic gearbox and cruise control.

2.2. Interfaces and Embedded Systems

We selected two vehicles with similar types of driving assistance (ACC and LCA) but different interfaces. The interface in the first car provided an exocentric representation of the road scene and a multimodal display interface. It displayed relevant information for driving a partially automated vehicle and took the limitations of the human cognitive system into account (Blömacher et al., 2020; Forster et al., 2019; Strand et al., 2018). This interface is therefore referred to hereafter as the *driving-centered interface*. The interface in the second car only provided visual information about the state of the automated systems, and is therefore referred to hereafter as the *vehicle-centered interface*. These interfaces differed on three major points: 1) the representation of the information regarding driving assistance; 2) the location of this information; and 3) the modalities used. The driving-centered interface represented information in an exocentric form (Tesla Model S, software version 8.0). The detection of road markings and other vehicles was indicated on the cluster. Icons representing the activation of Level 1 or Level 2 automation were displayed at the top of the cluster. The visual feedback was supplemented with auditory signals. When the Level 2 automated system (ACC + LCA) was activated or suspended, a sound was emitted. The sound of activation was a two-note rising tone. The sound of suspension or deactivation was a two-note falling tone. By contrast, the vehicle-centered interface only provided visual feedback (Volvo XC60). Pictograms representing the activation of automation were located on the cluster and the HUD. The steering wheel was slightly stiffer at Level 2 than at Level 1 for both vehicles, but the strength needed to override the lateral control was slightly greater for the driving-centered interface than for the vehicle-centered interface. A description of the technical features of the two interfaces is provided in Table 8. Figure 11 represents the instrument's cluster and the Level-2 activation's icon. Both vehicles required specific conditions to activate their automated system. Two clear road markings were necessary for the activation of Level 2. The driving-centered interface indicated the availability of the automation by highlighting the road marking representation in grey on the cluster. If the system could no longer detect one of the road markings, it switched itself off. The main reason for suspension was failure to detect a road marking. The one major difference between the two types of automation was that the Tesla was able to perform automatic lane changes. When Level 2 was activated, activation of the turn signals automatically resulted in a lane change.

Table 8

Display of onboard system activation by the driving-centered interface and vehicle-centered interface. The HUD of the vehicle-centered interface is shown on the righthand side.

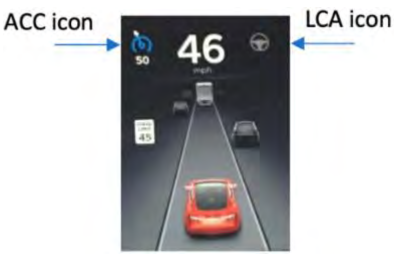
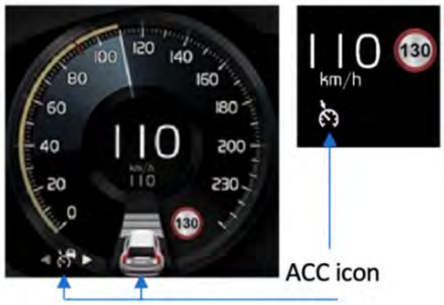
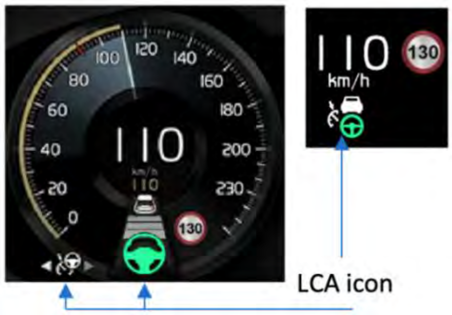
		Driver-centered interface	Vehicle-centered interface
No assisted systems			
Display of activated systems	ACC (Level 1)		
	ACC + LCA (Level 2)		
Activation commands	ACC (Level 1)		
	ACC + LCA (Level 2)		

Figure 11

Photos showing the instrument's cluster and forward views of each vehicle with either the driving-centered interface (left) or vehicle-centered interface (right).

The letter A indicates the icon displayed when LCA and ACC are activated.

2.3. Task

Participants had to drive the vehicle they were assigned on a pre-established circuit. They were instructed to drive normally, as though they were driving their own car. They were encouraged to use the automated system whenever they wanted and whenever they felt safe. They were instructed to respect the speed limits and traffic laws. A satnav system indicated the directions both visually (on the central multimedia screen) and verbally in each vehicle. During the drive, the experimenter asked the driver questions. The questions were asked when the driver was cognitively available (e.g., while driving along a straight road), so as not to increase cognitive load. Drivers had to respond as spontaneously and sincerely as possible. These questions were designed to evaluate the drivers' understanding of the automation's state and functioning.

2.4. Road circuit

The road circuit was a 45-minute round trip along public roads and highways, at an average speed of 75.75 km/hr. First, participants drove along a divided highway for 4 km. Then came a 9-km stretch of motorway (four lanes of traffic travelling in each direction for 4 km, then three lanes in each direction for 5 km), followed by a further 11 km on a divided highway. They then turned round at a roundabout, and followed the same itinerary in the opposite direction until they reached their starting point. For the entire round trip, the two sides of the road were separated from each other by a central reservation, and the road was mostly straight or slightly curved.

2.5. On-road Questioning

Participants were asked on-road questions to assess their awareness of the functioning and level of automation of the vehicle they were driving. To answer the questions, drivers either had to look for the requested information in their environment or else retrieve it from memory. The questions covered four topics. The first type of question (*system questions*) was system-oriented and concerned the status and functioning of the vehicle (e.g. "What is your current speed?" or "Is the ACC activated?"). The second type of question (*personal questions*) covered personal information about the driver (e.g. "Do you have children?"). The third type of question (*interior questions*) covered information accessible inside the car (e.g. "What radio station are we listening to?"). The last type of question (*exterior questions*) covered information accessible outside the car (e.g. "Is the car behind us black?") (see Appendix A for the list of questions). The system questions were mixed with the other types of questions, so that participants were not able to prepare their answers. The order of the questions was randomized. The answers to the questions were recorded. The start and end times of each question and each answer were extracted with Audacity software for the vehicle-centered interface group, and BeGaze software for the driving-centered interface group.

2.6. Eye tracker

Our choice of eye-tracking technique took into consideration the areas fixated by drivers during the experiment while answering the questions. This measure allowed us to ensure that drivers knew where to find the required information. We used SensoMotoric Instruments (SMI) Eye Tracking Glasses, which are equipped with infrared sensors to monitor eye movements (saccades, fixations and blinks) and a front camera to record the field of vision. The eye-tracking data were recorded at a sampling frequency of 60 Hz. The glasses were connected to a mobile phone (Samsung Galaxy Note 4) that allowed us to power the glasses, calibrate the gaze measures, monitor the visual behaviour in real time, and store the video and audio recordings. Eye-tracking data were extracted and processed using BeGaze 3.7 software. We also used this software to map the fixations. This mapping consisted in associating each recorded fixation with an area of interest (AOI). BeGaze software then calculated the gaze count and duration for each AOI. There is no explicit definition of a gaze in the BeGaze software, but we worked on the principle that a gaze is a fixation on an AOI if it remains within a 100-pixel area for around 20 ms.

2.7. Drivers' Comments

During the tests, participants were free to comment on their driving experience. The comments made during the entire experiment were recorded with the eye-tracker's microphone. As the microphone was placed on the bridge of the glasses not far from the mouth, the audio recording was good enough to hear the participant's comments. These comments were transcribed and classified. This method served to complement the eye-tracking and on-road questioning measures. The goal of the analysis was to identify mode confusions, which was not possible with the quantitative data. It also served to collect information about which aspects of the interface allowed drivers to understand the automated systems.

2.8. Procedure

Participants began the experiment by filling out a questionnaire. The training phase then began. During this phase, the basic controls were explained to the participants, as were specific aspects of the activation and suspension of the automated system. The driver's responsibilities regarding the use of each assistive device were explained to participants. Afterwards, the participants were equipped with the eye-tracking glasses. The volunteers then drove the vehicle for about 10 minutes on a straight road. Once the experimenter was sure that the volunteers were able to control the car and its features, the experimental phase began. Participants drove for about 45 minutes, following the pre-established circuit on the satnav. The experimenter sat in the front passenger seat and asked the participant questions without disturbing him or her. Finally, the experiment ended with a questionnaire and an interview.

2.9. Experimental Design

A 2 (between-participants) x 4 (within-participants) experimental mixed design was used. The first factor we manipulated was the driven vehicle, and it had two modalities (driving-centered interface vs. vehicle-centered interface). The second manipulated factor was type of question, and it had four modalities (system questions vs. personal questions vs. interior questions vs. exterior questions).

2.10. Measures & analysis

2.10.1. Response Times to On-Road Questions

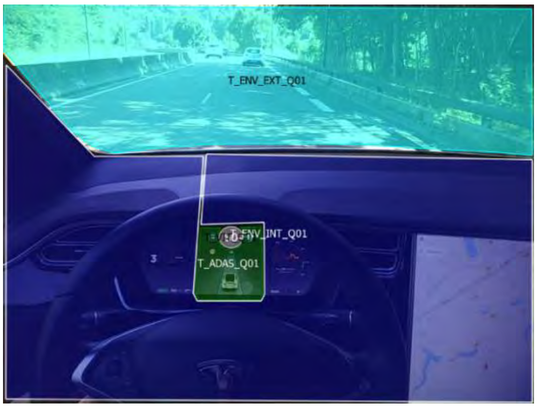
For the audio recordings of the drivers answering the on-road questions, we began by extracting the start and end times of each question and each answer. To measure the time needed by participants to access the information and verbalize it, we calculated the interval between the end of the question and the end of the answer. The assumption of normality of residuals for mean response times was not met for the driving-centered interface group. We ran a mixed variance analysis on means of median response times, with type of questions as a within-participants factor, and interface type as a between-participants factor. To investigate interface use in greater depth, the questions were analysed separately for each system. Given the small sample, we ran a two-tailed Mann-Whitney *U* test for each system question, with interface type as a between-participants factor. The Mann-Whitney *U* test is a nonparametric test that allows two small independent samples to be compared, in order to decide whether they are equally distributed or not (Mann & Whitney, 1947).

2.10.2. Fixation Count and Duration

Gaze positions were coded on a reference image featuring all the AOIs. We used four or five AOIs for each vehicle. For the driving-centered interface group, the AOIs were the external environment, the interior environment, the speed display, and the ADAS. The vehicle-centered interface had the same AOIs, plus the HUD (see Figure 12). For each AOI, we extracted the mean fixation duration and number of fixations during an 8-second window, extending from 4 s before to 4 s after the end of the question. This window allowed us to consider the processing time (i.e. listening to the question, finding the answer, and verbalizing it). The analyses compared the driving-centered interface group and the vehicle-centered interface group on the mean duration and number of fixations. Owing to the small sample and the non-respect of the normality of residuals for the driving-centered interface group, we ran nonparametric analyses. Nonparametric Wilcoxon rank-sum tests were carried out, with group as a between-participants factor for both measures. The same analyses were carried out separately for each system question on each AOI, except for the HUD. No participants were excluded from the analysis ($N = 20$).

Figure 12

AOIs used in the driving-centered interface (left) and vehicle-centered interface (right).



The HUD displayed the vehicle's speed and information about the automated system (see Table 1). The HUD AOI was too small for us to know for certain whether the fixations were on the speed information or on the automation information. We therefore did not include the HUD in the comparison between the two interfaces. We can, however, assume that during the questions about speed or the automated system, drivers looked at the HUD to obtain the relevant information. To consider the fixations on the HUD, we calculated descriptive statistics. For the vehicle-centered interface group, we calculated the means and standard errors of fixation count and duration on the HUD, speed display, and ADAS AOIs. The same operation was repeated for the driving-centered interface group, minus the HUD. We then calculated the mean fixation count and duration for each AOI for the questions on speed. We repeated the same operation for the questions on the automated system.

2.10.3. Classification of Drivers' Comments

The drivers' comments about the interface in relation to the automated system were transcribed and categorized, using a bottom-up approach. Three categories and 15 subcategories emerged. The three categories were (1) activation of ACC (Level 1) or ACC combined with LCA (Level

2), (2) deactivation of ACC (Level 1) or ACC combined with LCA (Level 2), and 3) mode confusions. Each category contained five subcategories. These subcategories corresponded to the communication medium between driver and system: (1) visual information display on the cluster; (2) auditory signals; (3) haptic feedback; (4) kinesthetic sensations; (5) commands; and (6) HUD (only for the vehicle-centered interface group). The analysis consisted in counting the number of comments in each category and subcategory.

3. Results

After indicating the response times for each type of question and each vehicle, we compare the vehicles on the system questions. Next, we describe the analysis of the eye-tracking data, comparing the two groups on fixation count and duration for the system questions. Finally, we describe the classification of the drivers' comments for each vehicle.

3.1. Response Times

We found a difference in the means of median response times between the groups and types of question (see Table 2). The ANOVA revealed an effect of type of question on response times, $F(3, 16) = 7.82$, $p < .001$, indicating that participants needed different amounts of time to answer the system questions, personal questions, interior questions, and exterior questions. Less time was needed to answer the system questions than either the personal questions, interior questions, or exterior questions. Regarding the effect of type of interface on response times for each type of question, the mixed ANOVA revealed no significant effect, $F(3, 16) = 3.45$, $p = .080$, $d = .08$. This means that participants took more or less the same time to answer each type of question regardless of which group they belonged to (see Figure 13). No interaction effect was observed between type of question and group ($p > .1$). For the system questions, groups differed slightly on the means of median response times (see Table 9). However, this difference was not significant ($p > .1$). Regarding the personal questions, we found a slight difference between the groups, but tests indicated that it was not significant ($p > .1$). This suggests that personal questions were not influenced by type of interface. For the interior questions, there was a small difference in mean response times, as well as in standard deviations. However, tests indicated that these differences were not significant ($p > .1$). For the exterior questions, we found a substantial difference. Driving-centered interface drivers seemed to answer questions about the external environment quicker than vehicle-centered interface drivers did. Again, however, tests indicated that this difference was not significant ($p > .1$). We then compared the two groups on response times for each system question.

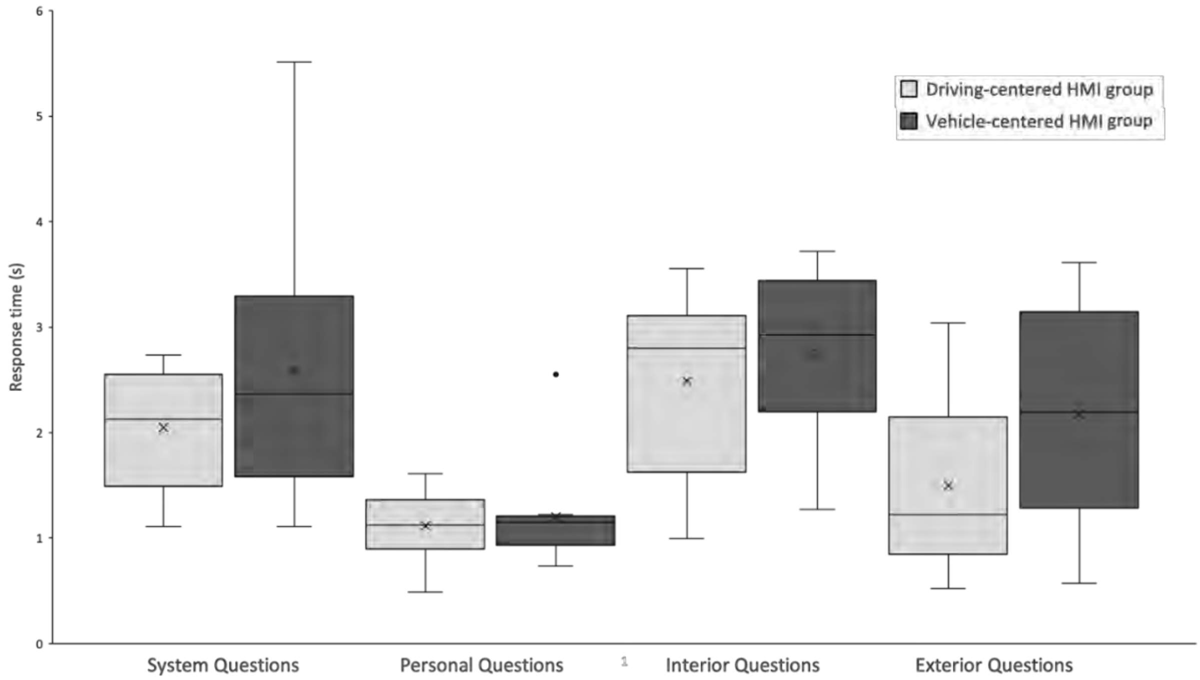
Table 9

Mean (standard deviation) response times depending on interface and question type.

	System Questions	Personal Questions	Interior Questions	Exterior Questions
Driving-centered interface group	1.61 s (0.46)	1.17 s (1.08)	2.27 s (0.84)	1.50 s (0.86)
Vehicle-centered interface group	2.16 s (1.50)	1.13 s (0.41)	2.55 s (1.04)	2.17 s (0.74)

Figure 13

Boxplot of response times depending on type of question and interface group.



4.2. System questions

The Mann-Whitney tests on each system question revealed significant differences between the two groups (see Figure 14). For the question "What is your current speed?" when the automated system was active, analysis revealed a significant difference between the groups, $U = 14$, $p = .02$, $r = -0.55$. On average, driving-centered interface users gave their speeds faster ($Mdn = 1.15$) than vehicle-centered interface users did ($Mdn = 1.82$). For the question "What is your current speed?" when the automated system was inactive, analysis revealed no significant effect, $U = 6$, $p = .07$, $r = -0.51$. Driving-centered interface users responded slightly faster ($Mdn = 1.43$) than vehicle-centered interface users did ($Mdn = 2.64$). Analysis of responses to the other system questions failed to reveal any significant differences ($p > .1$; Table 10).

Figure 14

Boxplot of response times depending on system question and interface group

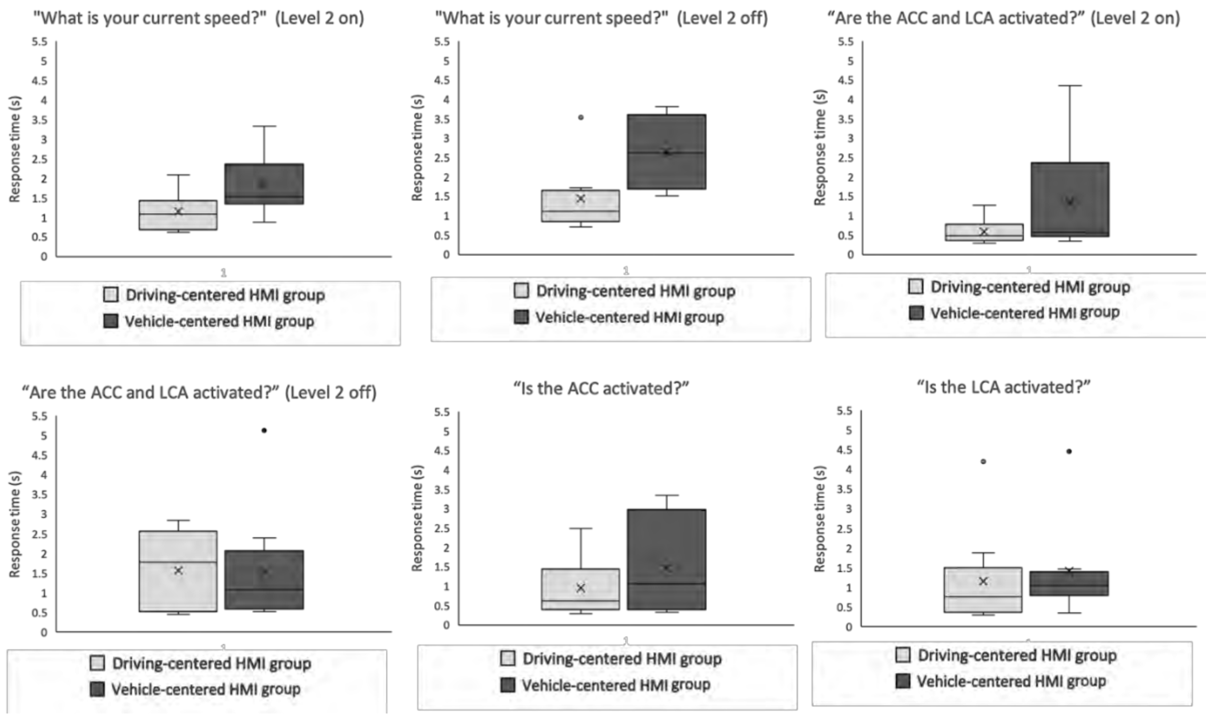


Table 10

Sample sizes and median (interquartile range) response times depending on the question and the interface.

** $p < .05$.*

System questions	Driving-centered interface group	Vehicle-centered interface group	Test results
"What is your current speed?" (Level 2 on)	(n = 10) 1.08 s (0.74)	(n = 8) 1.53 s (1.01) *	U = 14, p = .02, r = -.55
"What is your current speed?" (Level 2 off)	(n = 9) 1.11 s (0.08)	(n = 4) 2.61 s (1.91)	U = 6, p = .07, r = -.51
"Are the ACC and LCA activated?" (Level 2 on)	(n = 10) 0.49 s (0.41)	(n = 9) 0.56 s (1.90)	U = 30, p = .24, r = -.28
"Are the ACC and LCA activated?" (Level 2 off)	(n = 9) 1.78 s (2.05)	(n = 9) 1.09 s (1.49)	U = 43, p = .86, r = -.05
"Is the ACC activated?"	(n = 10) 0.62 s (1.05)	(n = 10) 1.06 s (2.59)	U = 36, p = .32, r = -.24
"Is the LCA activated?"	(n = 9) 0.76 s (1.14)	(n = 8) 1.05 s (0.60)	U = 25, p = .32, r = -.26

3.2. Eye tracking

The subsequent analyses served to compare the two groups on each measure, each AOI, and each system question. Because of the large number of analyses (more than 50), only the significant effects and those tending toward significance are reported here. For the question "What is your current speed?" when the ACC and LCA were active, no significant difference was observed between the groups on the mean fixation time on the speedometer, $U = 17$, $p = .06$, $r = -0.449$ (see Table 11). Driving-centered interface users made slightly longer fixations on the speedometer (*Mdn* = 294 ms) than vehicle-centered interface users did (*Mdn* = 75 ms; see Fig. 5). For the question "What is your current speed?" when in manual driving mode, analysis revealed an effect of group on the number of fixations on the speedometer, $U = 3$, $p = .02$, $r = -0.656$. Driving-centered interface users looked at the speedometer more often (*Mdn* = 2) than vehicle-centered interface users did (*Mdn* = 0). Groups did not differ significantly on mean fixation time for this question, $U = 5.5$, $p = 07$, $r = -0.524$. On average, driving-centered interface users made slightly longer fixations (*Mdn* = 523 ms) on the speedometer than vehicle-centered interface users did (*Mdn* = 0 ms). For the question "Is the ACC activated?" when it was activated, the test revealed no significant effect on the number of fixations on the ADAS AOI, $U = 20.5$, $p = .07$, $r = -0.422$. Driving-centered interface users tended to look slightly more often (*Mdn* = 6) at the ADAS than vehicle-centered interface users did (*Mdn* = 0). A significant difference was observed for the number of fixations on the interior of the vehicle for the same question, $U = 17$, $p = .03$, $r = -0.422$. Vehicle-centered interface users (*Mdn* = 1) made more fixations on the interior of the vehicle than driving-centered interface users did (*Mdn* = 0). In addition, there was a significant difference between the groups on mean fixation duration outside the vehicle for the same question, $U = 11$, $p = .009$, $r = -0.614$. Vehicle-centered interface users (*Mdn* = 5793 ms) looked outside the vehicle for significantly longer time periods than the driving-centered interface users did (*Mdn* = 4066 ms; see Figure 15 for a graphical representation).

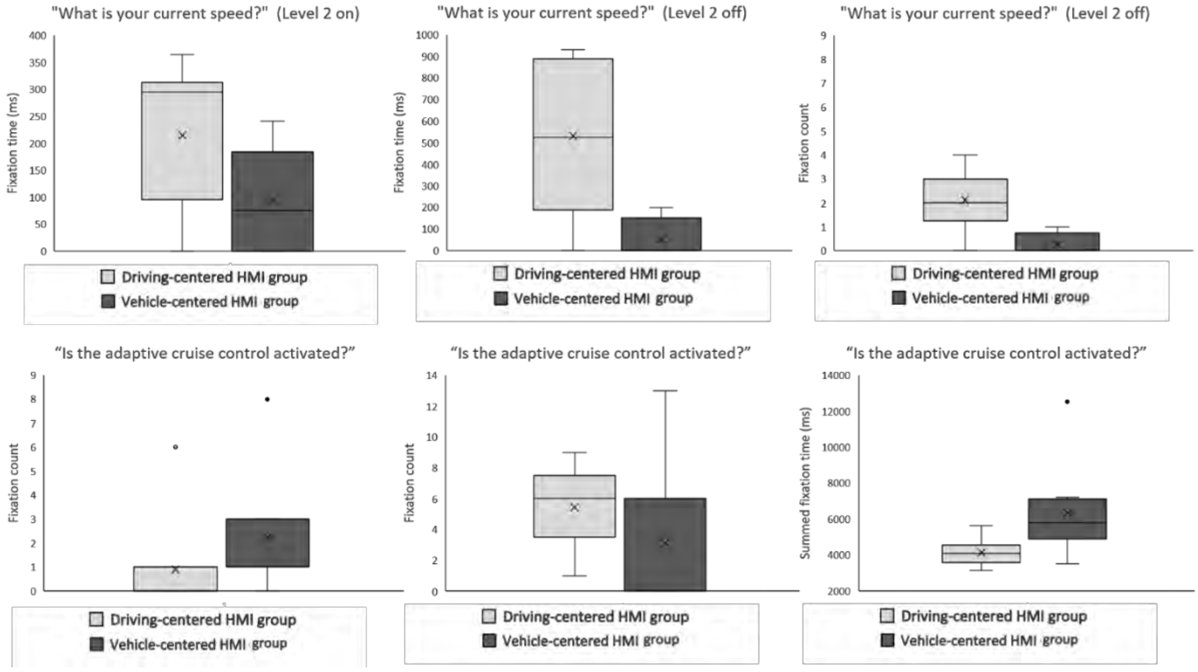
Table 11

Sample sizes, median (interquartile range) numbers and mean durations of fixations depending on the vehicle, and results of nonparametric analysis.
*Significant differences are in italics. * p < .05. ** p < .01.*

Question	AOI	Measure	Driving-centered interface	Vehicle-centered interface	Test results
"What is your current speed?" (Level 2 on)	Speedometer	Fixation duration	(n = 9) 294 ms (194)	(n = 8) 75 ms (171)	U = 17, p = .06, r = -0.45
"What is your current speed?" (Level 2 off)	Speedometer	Number of fixations	(n = 8) 2 (2.5) *	(n = 4) 0 (1.25)	U = 3, p = .02, r = -0.66
		Fixation duration	(n = 8) 211.6 ms (150)	(n = 4) 0 ms (171)	U = 5.5, p = 07, r = -0.52
"Is the adaptive cruise control activated?"	ADAS	Number of fixations	(n = 9) 6 (3)	(n = 9) 0 (4)	U = 20.5, p = .07, r = -0.42
	Interior of the vehicle	Number of fixations	(n = 9) 0 (3) *	(n = 9) 1 (4)	U = 17, p = .03, r = -0.42
	External environment	Summed fixation duration	(n = 9) 4066 ms (63) **	(n = 9) 5793 ms (114)	U = 11, p = .009, r = -0.61

Figure 15

Boxplots of fixation counts and durations depending on the question and the interface group.



For the vehicle-centered interface group, the number of fixations on the HUD and their mean duration were non-negligible (see Table 12). The number and duration of fixations on the HUD were the same as for the speedometer (for speed questions) and the ADAS (for automated system questions). This suggests that the HUD was used just as much as the speedometer and the ADAS to answer system questions.

Table 12

Mean (standard deviation) fixation count and duration on the HUD, speedometer and ADAS. AOIs of the vehicle-centered interface and driving-centered interface groups for each type of question.

Question	AOI	Vehicle-centered interface		Driving-centered interface	
		Fixation count	Mean fixation duration	Fixation count	Mean fixation duration
Speed questions	Speedometer	3.20 (2.95)	146 ms (71)	2.40 (1.06)	227 ms (118)
	HUD	4.25 (2.25)	280 ms (117)	-	-
System questions	ADAS	4.91 (3.61)	201 ms (124)	2.36 (1.47)	183 ms (103)
	HUD	3.20 (1.67)	336 ms (289)	-	-

3.3. Drivers' comments

Every participant made spontaneous comments about the driving experience, interface, and functioning of the automated system. These comments are described separately for each vehicle. We divided them into three categories: (1) activation of automated system, (2) deactivation of automated system, and (3) mode confusions. We counted the number of comments in each category. Most of the vehicle-centered interface users referred to the cluster to comment on the activation of the automated system (see Table 13). Four of the 10 vehicle-centered interface users referred to the HUD. One vehicle-centered interface user referred to the commands, and four to the haptic feedback. Finally, four participants in the vehicle-centered interface group reported relying on kinesthetic sensations. For the deactivation of the automated system, three of the 10 vehicle-centered interface users mentioned the commands. Only one vehicle-centered interface user reported using the cluster. The most frequently encountered mode confusions for the vehicle-centered interface drivers were between Levels 1 and 2. Three vehicle-centered interface users became confused because of unclear information on the cluster. One vehicle-centered interface user reported uncertainty or mode confusion between manual driving and ACC (Level 1) when consulting the visual information on the cluster. Another user reported mode confusion between manual driving and the ACC (Level 1) when using the commands.

Table 13

Number of participants in the vehicle-centered interface group (total number of participants in the group) who commented about a specific modality of the interface regarding activation, deactivation, or mode confusions.

	Activation of automated system	Deactivation of automated system	Mode confusion between manual driving and Level 1 systems	Mode confusion between Level 1 and Level 2 systems
Visual	8 (10)	1 (10)	1 (10)	3 (10)
Commands	1 (10)	3 (10)	1 (10)	0
Haptic feedback	4 (10)	0	0	0
Kinesthetic sensations	4 (10)	0	0	0
HUD	4 (10)	0	0	0

Five of the 10 users of the driving-centered interface reported relying on the visual modality of the cluster to know if the automation was activated (see Table 14). Only one participant reported using the auditory feedback. Two participants referred to the commands. One participant referred to the haptic feedback, while another referred to kinesthetic sensations. As for deactivation of the automated system, two of the 10 driving-centered interface users relied on visual information displayed on the cluster, two referred to the commands, two to the auditory feedback, and one to the haptic feedback. Three participants experienced confusion between Level 2 (ACC + LCA) and either Level 1 (ACC) or manual driving when consulting the visual information. The same kind of confusion was reported by three participants when listening to the auditory feedback. Confusion between Level 1 and either the Level 2 or manual driving was experienced by one driving-centered interface user when consulting the visual information on the cluster. Confusion between Level 1 and Level 2 was experienced by one participant in the driving-centered interface group when consulting the visual information. The same kind of confusion was reported by one driving-centered interface user who relied on kinesthetic sensations.

Table 14

Number of participants in the driving-centered interface group (total number of participants in the group) who commented about a specific modality of the interface regarding activation, deactivation, or mode confusions.

	Activation of automated systems	Deactivation of automated systems	Mode confusion between Level 2 and Level 1 or manual driving	Mode confusion between Level 1 and Level 2 or manual driving	Mode confusion between Levels 1 and 2
Visual	5 (10)	2 (10)	3 (10)	1 (10)	1 (10)
Auditory	1 (10)	2 (10)	3 (10)	0	0
Commands	2 (10)	2 (10)	0	0	0
Haptic feedback	2 (10)	1 (10)	0	0	0
Kinesthetic sensations	1 (10)	0	0	0	1 (10)

4. Discussion

Our study was designed to explore the impact of interface design on the detection of information about the status and functioning of two partially automated vehicles. We compared two interfaces: a driving-centered interface and a vehicle-centered interface. The driving-centered interface used an exocentric representation of the road scene and multimodal information to inform drivers about the level of automation. The vehicle-centered interface used a traditional display with visual information only. During a 45-minute on-road driving session, response times and eye movements were recorded while participants answered questions. Spontaneous comments were also recorded, and subsequently classified according to whether they were related to the interface or the state of the automated systems. In this section, we discuss the extent to which our results answer the question of whether the exocentric and multimodal driving-centered interface allows drivers to have a better understanding of the functioning of the vehicle when using an automated system than a classic visual vehicle-centered interface.

Overall, response times for system questions were shorter for the driving-centered interface group than for the vehicle-centered interface group. More specifically, the two groups differed on response times for questions about vehicle speed. When the automation was on, driving-centered interface users responded faster than vehicle-centered interface users. These results suggest that vehicle speed information is more accessible in the driving-centered interface than in the vehicle-centered interface. However, we failed to find any differences in response times to questions about the activation of ACC (Level 1) or ACC + LCA (Level 2). Compared with the classic interface, the multimodal and realistic driving-centered interface seemed to inform drivers more efficiently about the functioning of the vehicle, though not about the level of automation. Regarding the eye-tracking data, for the speed questions, the driving-centered interface group made more/longer fixations than the vehicle-centered interface group on the speedometer. At first glance, our results suggest that the exocentric interface captured the drivers' visual resources more than the classic interface did when it came to vehicle speed. This suggests that it may have been difficult to extract visual information (Jacob & Karn, 2003). However, we observed a high fixation count and duration on the HUD, such that the vehicle-centered interface drivers fixated the HUD at least as many times and for at least as long as they did either the speedometer or the ADAS. The fact that participants looked at both the speedometer and the HUD when answering the speed question suggests that they double-checked. They may have seen the information on the HUD and verified it on the speedometer. The time needed to perform this double checking may explain why it took the vehicle-centered interface users longer to answer questions about vehicle speed. The use of the HUD in this situation may have been more demanding than that of the driving-centered interface. One way to measure the benefit of the HUD would be to compare the answers of vehicle-centered interface drivers to speed questions in two HUD conditions: activated versus suspended. This would make it possible to evaluate the HUD's impact on response times.

Regarding ACC activation (Level 1), results diverged. Driving-centered interface users looked more often at the ADAS, where the information was displayed. By contrast, vehicle-centered interface users looked more often at the interior of the vehicle, suggesting that they were looking for the information in the wrong places. However, they also looked at the external environment for longer periods than the driving-centered interface users, which suggests better situational awareness. We had expected driving-centered interface users to make shorter and fewer fixations on the cluster than the vehicle-centered interface users when answering questions on automation. The multimodal driving-centered interface provided auditory feedback about the level of automation, meaning that users could access the information without having to look at the cluster. However, we found no significant differences in fixations compared with the classic interface. This suggests that the multimodal interface did not correctly inform drivers about the state and functioning of the vehicle. This result was confirmed by the mode confusions encountered by the driving-centered interface group.

We observed a wider range of mode confusion situations for the driving-centered interface group than for the vehicle-centered interface group (see Tables 6 and 7). For the driving-centered interface users, most of the confusions were related to Level 2 (ACC + LCA), and involved both visual and auditory information. Thus, instead of helping drivers to distinguish between the modes, the multimodal interface actually induced confusion in some of them. This may be an indication that the auditory information was not correctly perceived or understood by users. In the vehicle-centered interface group, most of the mode confusions concerned Levels 1 (ACC) and 2 (ACC + LCA). There was only a subtle distinction between these two levels on the cluster, which may explain these confusions. Fewer confusions related to the ACC were observed for the driving-centered interface. The use of a realistic representation of the vehicle being followed may be an efficient clue, helping drivers to understand the activation state of the ACC (Level 1).

According to the multiple resources model, using different sensory channels to convey information is more efficient than using a single one (Wickens, 2008). According to the automated driving model, an interface that informs drivers about the tasks undertaken by the automated systems is more efficient (Stanton et al., 2001). On the basis of these two models, we had assumed that the multimodal and exocentric driving-centered interface would allow drivers to understand the vehicle's status and capabilities more efficiently than the vehicle-centered interface would. Despite the presence of an HUD in the vehicle-centered interface, the driving-centered interface allowed drivers to respond faster to questions about the vehicle's level of automation. Interestingly, however, the driving-centered interface captured more visual attention overall than the vehicle-centered interface did. This result is consistent with the results of Gaspar and Carney (2019). In their study, participants used a Tesla Model S (driving-centered interface) to do daily drives, and their eye movements were recorded. The authors found that users looked at the external environment less frequently when Level 2 was activated. Our results corroborate these findings, suggesting that an exocentric representation tends to be visually distracting. Regarding mental models, the exocentric and multimodal driving-centered interface seemed to induce mode confusions. Banks et al. (2018) observed similar results with participants using a Tesla Model S with an interface similar to the one we used in our study. These authors reported mode confusions with the activation of Level 2 automation. Mode confusion with the activation of Level 2, specifically relating to visual information, was also reported in the present study. Interestingly, however, we observed more mode confusion with the deactivation of Level 2. This confusion was related to the auditory feedback intended to inform drivers about the activation or deactivation of the automated system. Participants' verbal comments indicate that some of these sounds were not understood. Carsten and Martens (2019) suggested that an efficient automated vehicle interface design implies that the functioning of the vehicle can be understood. If the meaning of signals from the interface is not understood, an accurate mental model cannot be forged. This is particularly important when the signal serves to inform the driver of a return to manual control. Many sounds are emitted nowadays in cars. Most of them are meant to alert drivers of an emergency. Designers should focus on sounds that efficiently inform them about the activation level of the automated system.

4.1. Conclusion

The purpose of our study was to compare the characteristics and on-road use of two interfaces. More specifically, our protocol was intended to determine which of the two designs best helped drivers understand the status and functioning of their vehicle. The first interface was multimodal and featured an exocentric representation of the road scene. The second interface was more traditional and used codes that are widely used in the automobile industry. The two designs gave rise to similar performances on questions about the status and functioning of the vehicle. However, the driving-centered interface was more efficient when it came to informing drivers about current speed. Eye movements suggested that its multimodal and realistic representation was more visually demanding. Finally, the driving-centered interface induced mode confusions related to Level 2, whereas the classic interface induced mode confusions related to Level 1. Taken together, results indicate that the multimodal driving-centered interface did not help participants driving the vehicle for the first time to gain a better understanding of the functionalities and activation level of the automated system. The mode confusions and failure to use the auditory information suggest that the sounds indicating activation and deactivation of automation could be improved. Future experiments should measure interface use over an extended period of time.

4.2. Limitations

On-road studies have just as many disadvantages as they do advantages. In ecological situations, some factors cannot be controlled. Variables dependent on the surrounding environment can interfere with the protocol. For example, the experiments with the vehicle-centered interface took place during vacation periods. The participants who drove with this interface were therefore more likely to encounter traffic jams. When in a traffic jam, the vehicle's speed is often too low to activate the automation. The vehicle-centered interface participants therefore used the automated system for shorter periods of time than the driving-centered interface users did. Another limitation arising from the ecological situation concerns the intrinsic differences between the vehicles. The main difference between the vehicles here was the presence of an HUD in the vehicle-centered interface but not in the driving-centered interface. Drivers could obtain several types of information from the HUD, as it centralized information about automation state and speed. Given the broad range of information and the relatively small surface area, we were not able to compare eye-tracking data for this AOI with that of the other AOIs. However, analysis of participants' comments suggested that the HUD was an important source of information on automation state.

Another limitation was the small sample size. With only 10 participants per group, caution needs to be exercised when considering the conclusions we drew. Except for one ANOVA, we only carried out nonparametric tests. Several differences were significant, but a larger sample size would have allowed us to increase the statistical power and strengthen the possible effects. This is one reason why we considered the drivers' comments. These supplemented the quantitative data by explaining what the participants experienced. Mode confusions are difficult to reproduce and observe in simulators. They can be accessible if drivers are interviewed, but this is generally done after the experiment, and only rarely while they are driving. Recording and transcribing drivers' comments allowed us to access mode confusions and explain the quantitative data. This exploratory study yielded relevant information about the impact an interface may have on drivers' awareness about mode state. However, it is important to note that this experiment only told us about what happens while drivers are familiarizing themselves with a vehicle. A 1-hour drive is not long enough for them to understand every feature of the interface, and familiarity with automated systems can have an impact on their utilization (Solís-Marcos & Kircher, 2019).

4.3. Future Research

As well as emphasizing the need for the functioning of the vehicle to be understood, Carsten and Martens (2019) underscored the need to carefully calibrate trust. Too much trust leads to disengagement, while too little trust discourages drivers from using the automated system. These authors argued that the exocentric representation of the driving-centered interface may simply serve to promote trust in the system. It may be esthetically appealing, but possibly not that useful in helping drivers take back control when necessary. It may be more suited to Level 3 vehicles, where the automated system gives drivers more time to take back control. In Level 2 vehicles, drivers need to be ready to do so at all times. The calibration of trust is therefore crucial. Future research should focus on looking for interfaces that foster the right degree of trust. One idea would be to indicate the vehicle’s limitations. Beller et al. (2013) explored the idea of warning drivers about the automation’s uncertainty in order to bring them back into the loop. In highly automated vehicles, the display features an icon representing uncertainty in situations where the automation may fail. In this condition, drivers are faster at taking back control and have greater trust in the automation. This kind of display could be included in Level 2 vehicles.

Another point to consider in future studies is familiarity with the system. Our study focused on the first driving session that participants had with the vehicle. A driver who is used to a Level 2 vehicle should know when the system is able to function properly or not. However, Solís-Marcos and Kircher (2019) observed a potential effect of familiarity with automation on eye movements. More specifically, the more experienced participants performed more secondary tasks, potentially leading to a loss of situational awareness. Future research should take experience with the system into account by featuring longitudinal experiments.

Points clés

- Des confusions de mode se sont produites dans le véhicule utilisant une interface multimodale centrée sur le conducteur et dans le véhicule utilisant une interface visuelle centrée sur le véhicule.
- Une interface multimodale centrée sur les besoins du conducteur permet de mieux comprendre le fonctionnement des systèmes automatisés, mais elle est visuellement distrayante.
- Les signaux auditifs de l'interface multimodale centrée sur les besoins du conducteur sont liés aux confusions de mode.
- L'indication d'informations sur la fiabilité pourrait être un moyen d'améliorer la conscience des modes.
- Des études longitudinales sont nécessaires pour étudier le processus d'apprentissage de l'utilisation des interfaces et des systèmes automatisés.

Key points

- Mode confusions occurred in vehicles using multimodal driver-centered interfaces and in vehicles using visual vehicle-centered interfaces
- A multimodal interface centered on the drivers' needs increases understanding of the automated system but is visually distracting.
- Auditory signals of the multimodal interface centered on the drivers' needs were related to mode confusions.
- Indication of reliability information could be a way to improve mode awareness.
- Longitudinal studies are necessary to investigate the process of learning to use interfaces and automated systems.

Chapter 5 – Driver compliance with automation reliability information regarding hazardous environmental circumstances

The present study's goal was to evaluate the utility of reliability information when environmental conditions varied favourably or unfavourably for automated systems. For that, we investigated the influence of reliability information and environmental conditions on drivers' judgment to the decision to deactivate the automated systems and take back control of the vehicle in partially automated cars. Thanks to a scenario-based method rooted in the Integrated Information Theory (Anderson, 2013), the interaction between the presentation of reliability information and environmental condition is studied and quantified. This study's results highlighted the utility of reliability displays. This experimental study was published in the journal Le Travail Humain and reformatted for the purpose of this manuscript. The formatting consisted in reducing the introduction section of the article to avoid redundancy with previous chapters.

Monsaingeon, N., Caroux, L., Langlois, S., Hurgobin, Y., & Lemercier, C. (2020). Driver compliance with automation reliability information regarding hazardous environmental circumstances. *Le Travail Humain*, 83(4), 343-360.

Résumé

Les systèmes automatisés des véhicules partiellement automatisés peuvent se suspendre soudainement et fréquemment en fonction des caractéristiques de l'environnement. Les constructeurs automobiles envisagent d'installer un indicateur visuel à l'intérieur du véhicule pour informer les conducteurs de la fiabilité des systèmes automatisés. La présente étude vise à évaluer l'influence des informations de fiabilité des systèmes automatisés et des conditions environnementales sur le jugement des conducteurs quant à la décision de désactiver les systèmes automatisés et de reprendre le contrôle du véhicule dans les voitures partiellement automatisées. Au total, 199 participants âgés de 19 à 67 ans ont été exposés à 16 scénarios réalistes décrivant des situations dans lesquelles un personnage conduit un véhicule partiellement automatisé. Ils devaient évaluer leur accord avec la décision du personnage de désactiver l'automatisation, en fonction de la couleur d'un indicateur d'approche des limites des systèmes automatisés (vert : fiable vs. orange : proche de la limite), du temps (très ensoleillé vs. pluie abondante), de la qualité du marquage routier (marquage clair vs. marquage flou), et de la courbure de la route (route droite vs. virages). Les résultats ont révélé que ces quatre facteurs ont influencé la décision de désactiver l'automatisation. Une analyse complémentaire par cluster a révélé que les résultats devaient également être interprétés en fonction du profil du conducteur (c'est-à-dire le sexe, la confiance dans les véhicules automatisés et l'expérience avec les véhicules automatisés). Ces résultats suggèrent que les profils des utilisateurs devraient être pris en compte pour décider d'inclure ou non un indicateur de fiabilité dans les véhicules partiellement automatisés.

Abstract

The automated systems of partially automated vehicles can suddenly and frequently deactivate depending on the characteristics of the environment. Manufacturers are considering the possibility of mounting a visual indicator inside the vehicle to inform drivers about the reliability of automated systems. The present study was designed to evaluate the influence of reliability information and environmental conditions on drivers' judgment to the decision to deactivate the automated systems and take back control of the vehicle in partially automated cars. A total of 199 participants aged 19-70 years were exposed to 16 realistic scenarios describing situations in which a character is driving a partially automated vehicle. They had to rate their judgment of agreement with the character's decision to deactivate the automation, depending on the color of an indicator of proximity to the limit of automation (green: reliable vs. orange: close to the limit), the weather (very sunny vs. raining heavily), the quality of the road markings (clear vs. blurred), and the curvature of the road (straight vs. bends). The results revealed that all four factors influenced the decision to deactivate the automation. A complementary cluster analysis revealed that the results also needed to be interpreted in relation to the driver's profile (i.e., gender, trust in automated vehicles, and experience with automated vehicles). These findings suggest that user profiles should be considered when deciding whether to include an indicator of reliability in partially automated vehicles.

1. Introduction

Partially automated vehicles have multiple limitations, which can cause the automation to suddenly deactivate. Researchers have been studying how best to display the limitations of vehicle automation for more than a decade (Seppelt & Lee, 2007). These displays have already been shown to be efficient in helping drivers to anticipate the system's limitations in poor weather conditions (Beller et al., 2013; Helldin et al., 2013). Drivers complied with the reliability display, meaning that they acted in accordance with the information. However, manual control recovery has been studied in relation with secondary tasks and training after the appearing of a failure (Payre et al., 2017), but not in relation with reliability displays before a potential suspension. How do drivers comply with reliability information in this context? Do they comply solely with the information about the reliability of the automation, or do they judge the situation according to all the environmental conditions? Moreover, how do experienced drivers, who already understand the system's limitations, react? Manufacturers are now considering installing reliability displays in partially autonomous vehicles, as evidenced by the launch of a European project in collaboration with manufacturers that aims to make automation safer[2]. However, the way in which these interfaces would impact the day-to-day use of automation, depending on the environment and the drivers' intrinsic traits, is not yet clear. Before overloading instrument clusters with yet more information, it is important to assess whether drivers would actually use it. In the present study, we evaluated how a reliability display is viewed, depending on different environmental circumstances. Drivers' characteristics, such as their trust in automation, experience with automated vehicles, and sex, were considered in relation to compliance with the display. After describing how automation works and its various limitations in the first two sections of the Introduction, we discuss previous studies of reliability displays. We then present a model predicting compliance with warning signals. Finally, we describe the scenario-based method we used to address our research question.

In all studies mentioned in automation, drivers appeared to take the reliability information into account, which impacted their driving performance. However, depending on the situation, drivers can choose whether or not to comply with (and act upon) the reliability information. In

[2] see for example the "ADAS & ME" project: https://www.adasandme.com/, retrieved on May, 12th, 2020.

Level-2 vehicles, drivers are faced with decisions. Depending on the situation, they can either deactivate the automation and take over control, or keep the automation activated. Reliability displays warn drivers of a probable suspension and necessary takeover. Noah and Walker (2017) suggested that the use of reliability displays should be interpreted in terms of compliance and reliance. *Compliance* is a positive response of the operator to a warning signal. *Reliance* is an absence of response by the operator to the absence of a warning signal. In the context of reliability displays for automotive automation, compliance is manifested by the operator's takeover when the display indicates high uncertainty. Given that indicators of limits of automation suggest a manual control recovery before a problem appears, compliance will be the focus of this study. Meyer (2004) included the reliance and compliance concepts in a model of decision-making in dynamic situations. This model predicts that compliance and reliance will be influenced by three main types of factor: normative, task, and operator. *Normative factors* relate to the situation and the ability of the warning signal to accurately signal a danger. Situation-related factors include the properties of the situations (e.g., risks caused by the weather). *Task factors* include the interface of the warning signal, the demand of the task, and its structure. Finally, *operator factors* incorporate general characteristics such as the skills of the operator and system-specific characteristics such as trust and belief about automation and prior experiences with the system. According to this model, compliance with warning signals is influenced by several factors. In the present study, we investigated the impact of normative and operator factors on compliance with reliability information, using an original experimental protocol regarding human-display interactions in cars.

As discussed above, system activation is sensitive to the quality of the relevant sensors. Several studies have explored situations where the weather deteriorates (Beller et al., 2013; Helldin et al., 2013), but none has so far focused on situations in which the road characteristics change (e.g., blurred road markings or a bend that is too sharp for the car's current speed; Endsley, 2017). Authors have yet to establish how drivers comply with reliability information when several environmental factors covary. To investigate how humans interact with an interface, researchers can use epidemiological data (accidentology), surveys (based on past or present experiences), on-road studies, and simulation studies. Simulations provide an opportunity to precisely investigate the interaction of a human with a display. However, they require substantial investment in both time and money. The scenario-based method developed by Anderson (2013) in his integrated information theory (IIT) may offer a viable alternative to regular methods. This IIT method allows several factors and their mutual interactions to be investigated at the same time. It relies on scenarios where participants are asked to evaluate combinations of factors, rather than individual ones. A full factorial design is required to determine the impact of each individual factor on overall judgments, and all possible interactions can be investigated (Anderson, 2013). This method has been validated in several different areas, including sport (Fruchart & Mullet, 2012), moral judgments (Trémolière & Djeriouat, 2016), and food purchase decisions (Hurgobin et al., 2019). Scenario-based studies have already been conducted in the automobile field (Bazilinskyy & de Winter, 2015), but not with the IIT method, despite its potential to explore new interfaces. The present study was thus the first to use the IIT method to evaluate interactions with a reliability display.

To understand how drivers comply with information about the limits of automation, environmental factors have to be considered. For example, the literature tells us that in poor weather conditions, reliability information can be beneficial. However, other conditions are also important, such as road curvature, as described by Endsley (2017), a factor that has not yet been studied but which manufacturers say would have an impact on the functioning of automation. The advantage of the IIT method is that it considers the judgments that people make in their daily lives. In the present study, the judgment regarded the decision to deactivate the automated systems. The first objective was to examine the interaction between situational factors that are liable to cause suspension (bad weather, sharp bends, blurred road markings) and compliance with the reliability information. According to Meyer's (2004) model, compliance with reliability information should be influenced by normative factors. The second goal was to determine the extent to which the operator's personal traits influence compliance. Again according to Meyer's model, compliance with reliability information should be influenced by personal factors such as trust, knowledge, and prior experience with automation.

2. Material And Method

2.1. Participants

Participants were unpaid French-speaking volunteers, recruited via online social networks and mailing invitation. Half of the participants were coworkers from Renault Technocentre or IRT SystemX. The other half were students from the University of Toulouse Jean-Jaures, acquaintances, or family members. Those participants who were recruited via social media belonged either to groups specializing in questionnaires or to student groups. The message to invite to participate to the study followed the guidelines of Dillman et al. (2014). The goal of the study was described as the evaluation of a display that aimed to make automated vehicles safer. The participants were told that in order to participate, they needed to hold a driving licence. They were informed that they would have to read scenarios and to answer questions. The name and contact of the responsible for the study was explicitly given. The sample was composed of 199 participants, including 101 women. Ages ranged from 19 to 67 years (M = 38.49, SD = 13). Mean length of driving experience was 19 years (SD = 13), and three quarters of the participants were regular drivers who drove at least 3 times a week (n = 155). Regarding automation, 160 participants had already used a speed regulator before, 77 had used an ACC (Level 1), and 27 had already used an ACC coupled with an LCA (Level 2). Of all the participants who started the experiment (n = 400), 50% completed the whole survey. The study was explained in the online survey and all participants then completed an informed consent form. Full anonymity was respected.

2.2. Material

2.2.1. Scenarios

The material consisted of 16 realistic scenarios, written in French, built following Anderson's methodology (2013). The scenarios were composed according to a four within-participants factor design: IPLA color (green: reliable vs. orange: not reliable) × Weather (very sunny vs. raining heavily) × Road marking quality (RMQ; clear vs. blurred) × Road curvature (RC; straight road vs. bends). The environmental factors were chosen according to some critical situations reported in car user manuals (in particular in that of the Renault Clio 5 2019). A first version on the study was presented to 8 coworkers of Renault Technocentre. The comments of the testers revealed that the material, instructions and the task were adapted to begin the study. According to Anderson's (2013) methodology, the persona used in the scenario was adapted to the gender and age of each respondent. The use of personas that resemble the participant allow them to project themselves into the situation and avoids common bias of rating methods (Anderson, 2013). Popular names in France were used. For women aged over 40 years, the name Marie was used. For women under 40, the persona was named Julie. For men over 40 years, the name Jean was used. For men under 40, the persona was named Julien.

After reading each scenario (e.g., "Julie is on the road, going on a vacation. Her vehicle is driving along a departmental road. The weather is **very sunny**. The road is **straight**. The road markings are **clear**. The IPLA is **green**", see Appendix B for all the stories), participants had to answer the question scenario "If you were in the same situation as Julie/Marie/Jean/Julien, to what extent would you agree with deactivating the system and going back to manual driving?" Participants had to judge indicate their agreement with the decision to deactivate the automation on a 20-points rating scale ranging from *Don't agree at all* on the left to *Totally agree* on the right. The points were not visible to the participant but were used in the analysis to quantify the judgment.

2.2.2. Instructions

In the instructions, participants were asked to read each scenario carefully and to use the entire rating scale. The vehicle was described as partially autonomous and therefore capable of automatically adapting its speed, lane position, and dealing with bends. However, the driver was described as wholly responsible for monitoring the road, and had to keep his or her hands on the steering wheel. The system was presented as having a limited capacity. In view of its limitations, the manufacturer had integrated an IPLA. This indicator had two colors: green (system functioning well) and orange (system approaching the limits of efficient functioning). When the system reached its limits, it suspended itself and the driver had to take back control. In every scenario, the automated systems were activated, and the driver had her/his hands on the steering wheel and monitored the road scene.

2.2.3. Procedure

A link to the online survey was sent by email or was accessible on the social media groups. Participants clicked on the link and completed the survey on their own, without an experimenter. The procedure followed Anderson's recommendations (2013). They began the experiment by providing their informed consent. They were then shown the instructions. An initial familiarization phase featured eight scenarios, including the most extreme ones, in order to induce a broad range of responses. The subsequent experimental phase was composed of the 16 scenarios. After each scenario, participants had to rate their degree of knowledge about automated driving on a 5-points scale ranging from 1 (*Poor*) to 5 (*Good*). Participants also had to rate their trust in automated vehicles in response to the question "To what extent do you agree with the following sentence: 'I totally trust in the automation of driving'?" on a 7-points scale ranging from 1 (*Totally disagree*) to 7 (*Totally agree*). Participants indicated whether they had any prior experience of automated systems (none, speed regulator, ACC, or ACC coupled with LCA). Participants were informed that the survey would last 15 minutes. The survey was implemented on Qualtrics and took less than 15 minutes to complete (M = 14.15, SD = 7.58).

3. Results

For each scenario, we calculated the mean judgment of agreement with the decision of deactivating the automated systems as a score ranging from 0 to 20. The higher the score, the greater the agreement to deactivate the automation. In accordance with Anderson's methodology (2013), we analysed the data by performing an analysis of variance (ANOVA). Given the multiplicity of comparisons, the significance threshold was set at $p < .001$, and the Bonferroni correction was used for post hoc tests. First, we conducted a within-participants ANOVA on the whole sample, to examine the main effects of the four factors and their possible interaction effects. Second, we performed a cluster analysis to capture the respondents' different profiles. We used a nonhierarchical centroid-based method (Euclidean distances) called *k*-means clustering, as recommended by Hofmans and Mullet (2013). This technique is less sensitive to outliers and uses all the datapoints. With this analysis, participants were divided into three groups, depending on their responses. Within each cluster, we conducted within-participants ANOVAs. Finally, using χ^2 tests, we tested the clusters for statistically significant sociodemographic differences.

3.1. Analysis Conducted on the Whole Sample

We ran a 2 (weather: sunny vs. raining heavily) × 2 (Road curvature: straight vs. bends) × 2 (road markings quality: clear vs. blurred) × 2 (IPLA: green vs. orange) ANOVA on the ratings provided in the experimental phase. The main effects of the four within-participants factors were all significant at $p < .001$ and had large effect sizes ($\eta^2p > .50$): IPLA, $F(198) = 240.31$, $p < .001$, $\eta^2p = .55$; weather, $F(198) = 273.03$, $p < .001$, $\eta^2p = .58$; road markings quality, $F(198) = 205.22$, $p < .001$, $\eta^2p = .51$, and road curvature, $F(198) = 255.32$, $p < .001$, $\eta^2p = .56$. Pairwise comparisons were conducted to examine differences between the levels of each within-participants factor (see Table 15). Participants preferred to deactivate the automation when the IPLA was orange than when it was green. Concerning the weather, they were more prone to deactivate the automation when it was raining heavily than when it was sunny. Participants were also more inclined to deactivate the automation when driving around bends than on straight roads. Finally, they were more likely to deactivate the automation when the road markings were blurred than when they were clear. Effect sizes (η^2_p) indicated that weather had the greatest impact on the decision to deactivate the automation, followed by the road curvature. The factor that had the third largest effect size was IPLA. Finally, the road markings quality was the last factor in terms of effect size.

Table 15

Mean ratings (standard deviations) and results of post hoc tests, reported for the whole sample, for IPLA, weather, road curvature, and road marking quality.

Factors	M (SD)	t	d
IPLA			
Green	9.51 (3.78)	-15.50***	-1.099
Orange	14.30 (2.73)		
Weather			
Sunny	10.02 (4.10)	-16.52***	-1.171
Raining heavily	13.79 (3.15)		
Road curvature			
Straight	10.05 (4.03)	-15.98***	-1.133
Bends	13.76 (3.27)		
Road Marking Quality			
Clear	10.44 (4.19)	-14.33***	-1.015
Blurred	13.37 (3.51)		

*** $p < .001$.

Three interactions were significant, starting with the Weather × IPLA interaction, $F(1, 198) = 26.75$, $p < .001$, $\eta^2p = .12$. The effect of weather was greater when the IPLA was orange rather than green. The IPLA × Road curvature interaction was also significant, $F(1, 198) = 18.57$, $p < .001$, $\eta^2p = .09$. The effect of road curvature was greater when the IPLA was orange rather than green. Finally, the weather × Road markings quality interaction was significant, $F(1, 198) = 14.02$, $p < .001$, $\eta^2p = .07$. The effect of road markings quality was greater when it was raining heavily than when it was sunny. Other interactions were not significant.

3.2. Cluster Analysis and Profile Definition

The cluster analysis yielded three-cluster and four-cluster solutions. The Elbow Method was followed to determine what solution was optimal (Kodinariya & Makwana, 2013). According to this method, the three-cluster solution offered the smallest within-cluster variability. We ran a series of Weather × Road curvature × Road markings quality × IPLA ANOVAs on the data from each cluster. The first cluster ($n = 61$) was called *Automation Skeptics*, as the IPLA factor was the most important and participants in this cluster had a high tendency to deactivate the automation. The second cluster was called *Compliant with IPLA* and contained the most participants ($n = 104$). The third cluster ($n = 34$) was called *Automation Enthusiasts*, as participants had a low tendency to deactivate the automation, even in the worst scenarios ($M = 12.85$, $SD = 6.41$). For each cluster, all four main within-participants factors were significant at $p < .001$. Results of the pairwise comparisons are summarized in Table 16.

Table 16

Mean ratings (standard deviations) reported by each clusters for IPLA, weather, road curvature, and road marking quality.

	Clusters		
	Automation skeptics (n = 61)	Compliant with IPLA (n = 104)	Automation enthusiasts (n = 34)
Factors	M (SD)	M (SD)	M (SD)
IPLA			
Green	15.05 (4.05)	7.80 (4.17)	4.84 (2.76)
Orange	18.19 (1.58)	14.41 (3.34)	6.98 (3.40)
Weather			
Sunny	15.07 (4.10)	8.95 (4.96)	4.25 (2.00)
Raining heavily	18.17 (1.51)	13.25 (4.27)	7.56 (3.40)
Road curvature			
Straight	15.12 (4.01)	8.94 (4.93)	4.35 (2.27)
Bends	18.12 (1.82)	13.26 (4.30)	7.47 (3.33)
Road Marking Quality			
Clear	15.56 (4.18)	9.41 (5.10)	4.40 (2.36)
Blurred	17.68 (2.13)	12.79 (4.58)	7.42 (3.32)

3.2.1. Automation Skeptics

The first cluster (n = 61) was called Automation Skeptics because of the participants' high tendency to deactivate the automation. According to the effect sizes, the order of importance of the factors for this cluster was (1) IPLA, $F(1, 60) = 84.93$, $p < .001$, $\eta^2p = .59$, (2) Road curvature, $F(1, 60) = 76.17$, $p < .001$, $\eta^2p = .56$, (3), weather, $F(1, 60) = 57.91$, $p < .001$, $\eta^2p = .49$, and (4) road markings quality, $F(1, 60) = 43.67$, $p < .001$, $\eta^2p = .42$. Analysis revealed several interactions between the factors for this cluster. IPLA interacted with weather, $F(1, 60) = 53.63$, $p < .001$, $\eta^2p = .47$. When the indicator was orange, participants had a high tendency to deactivate the automation, whether it was sunny or raining heavily. When it was sunny and the IPLA was green, participants had a low tendency to deactivate the automation. When the IPLA was green and it was raining, participants had a high tendency to deactivate the automation. The IPLA therefore appeared to have a strong impact when the weather was sunny. When it was raining, the IPLA had a smaller impact. The same logic applied to the other interaction effects. For the interaction between the IPLA and road curvature $F(1, 60) = 26.32$, $p < .001$, $\eta^2p = .31$, the IPLA had a considerable impact when the road was straight. For the interaction between IPLA and road markings quality $F(1, 60) = 13.69$, $p < .001$, $\eta^2p = .19$, the color of the IPLA had a greater impact when the road was straight than when there were bends. These results suggest that for this cluster, the IPLA was a decisive factor when deciding to maintain the automation. The majority of Automation Skeptics did not show trust in automation, as 33% declared that they rather disagreed with the statement “I totally trust in automated cars”. This cluster was mostly composed of women ($n = 41$) and its members had little experience with automated driving, as only 25% of them had used ACC (Level 1) automation before.

3.2.2. Compliant with IPLA

The second cluster was called *Compliant with IPLA* and had the largest number of participants ($n = 104$). All four factors had a significant effect on agreement with deactivating the automation. According to the effect sizes, the order of importance of the factors for this cluster was (1) IPLA, $F(1, 103) = 205.46$, $p < .001$, $\eta^2p = .67$, (2) weather, $F(1, 103) = 195.51$, $p < .001$, $\eta^2p = .66$, (3), road curvature $F(1, 103) = 162.12$, $p < .001$, $\eta^2p = .61$, and (4) road markings quality $F(1, 103) = 134.60$, $p < .001$, $\eta^2p = .57$. For this cluster, analysis failed to reveal any interaction effects. In this cluster, which included 50 women, 30% rather disagreed with the statement “I totality trust automated cars”, and 20% rather agreed with it. Respondents were mildly experienced, as 40% of them had used ACC before.

3.2.3. Automation Enthusiasts

The third cluster ($n = 34$) was called *Automation Enthusiasts*, as its members had a low tendency to deactivate the automation. In this cluster, all four factors had a significant effect on agreement with deactivating the automation. According to the effect sizes, the order of importance of the factors for this cluster was (1) road markings quality $F(1, 33) = 33.75$, p

$< .001$, $\eta^2p = .51$, (2) weather, $F(1, 33) = 34.22$, $p < .001$, $\eta^2p = .50$, (3) IPLA, $F(1, 33) = 17.02$, $p < .001$, $\eta^2p = .34$, and (4) road curvature $F(1, 33) = 28.84$, $p < .001$, $\eta^2p = .47$. No interaction effects were found in this cluster. This cluster contained participants with the highest trust in automation, as 35% of respondents rather agreed with the statement "I totally trust automated cars". Most of the respondents had experience with automation, as almost 55% had used ACC in the past and 20% had used Level-2 automation. This cluster was mostly composed of men (24).

Sociodemographic composition of the clusters

The chi-square contingency table revealed several links between the clusters and the demographic data. Sex proved to be significantly linked to cluster formation, $\chi^2(2, N = 199) = 13.11$, $p = .001$. Trust in automated cars was also related to cluster formation, $\chi^2(2, N = 199) = 36.26$, $p < .001$, as was previous experience with automated systems, $\chi^2(2, N = 199) = 13.15$, $p = .041$. Neither age nor previous knowledge about automation was related to cluster formation.

4. Discussion

In this study, we evaluated the impact of providing an IPLA on the decision to deactivate automation, in relation to several environmental conditions. We used scenarios in which the IPLA color varied, along with the weather, the road curvature and the road markings. The first goal was to evaluate whether situational factors influenced IPLA compliance. The second goal was to determine whether respondents' personal traits influenced IPLA compliance. Results showed that the weather, IPLA, road curvature, and road markings quality all influenced the decision to take back control of the driving. When the IPLA was orange, corresponding to a less reliable system, respondents tended to agree with deactivating the automation. By contrast, when the IPLA was green, they tended to disagree with deactivating the automation. Respondents also took environmental factors into consideration. When it was raining heavily, drivers tended to agree with deactivating the automation. When the road markings were blurred or there were bends in the road, participants also tended to agree with deactivating the automation. Cluster analysis highlighted three profiles of respondents. Respondents with the first profile (*Automation Skeptics*) did not trust automation and had little experience with it.

These respondents tended to agree with deactivating the automation more often than respondents in the other two clusters did. The IPLA information was a priority for them. The second profile *(Compliant with IPLA*) was more heterogeneous. It included people who trusted automation, but others who did not. They had little experience with automation. These respondents considered the IPLA information. Participants with the last profile (*Automation Enthusiasts*) preferred to keep the automation activated most of the time. These experienced respondents considered the environmental factors just as much as the IPLA. Taken together, results indicated that the IPLA's color was a decisive factor in the decision to deactivate the automation. Most of the participants complied with the information and rather agreed with deactivating the automation when it was suggested. The usefulness of this indicator of the reliability of automation was confirmed by our results, and it could potentially be accepted by users in the future. These results were in line with the predictions of Meyer's (2004) model of compliance and reliance. Normative factors influenced compliance with the reliability display. When the weather and road curvature varied, compliance with the IPLA varied. Thus, when weather or road conditions were poor, compliance with the IPLA was very high. This highlights the importance of having a reliability display that is coherent with what the driver can perceive.

The cluster analyses allowed us to distinguish between three types of respondents: *Automation* Skeptics, *Compliant with IPLA*, and *Automation enthusiasts*. These groups differed on sex ratio, initial trust in automated vehicles, and experience with automation. These factors correspond to Meyer's *operator factors* that influence compliance. Our results seem to confirm the prediction of his model. *Automation Skeptics* deactivated the automation more often than the other groups did. They did not appear to rely on the automation. The respondents were mainly women, had little trust in automation, and little experience with automated systems. They based their judgment on the decision to deactivate the automation mainly on the color of the IPLA. They complied with the reliability indicator. For users with this type of profile, information about the reliability of automation seems appropriate, and may allow them to improve their understanding of the function of automation. However, this cluster was most likely to be at risk, as participants were liable to base their decision almost entirely on the IPLA, without paying attention to other determining factors, such as road markings quality. Road markings quality was the last factor that was taken into account, even though it is the most recurrent cause of suspension in Level-2 vehicles. It is possible that the little experience that the members of this group had with automation impacted their perception of the gravity of the situation and thus their use of automation. This result further supports the idea that having a transparent interface allows users to understand the right conditions for using the automation. Participants in the largest cluster, *Compliant with IPLA*, also based their decision mostly on the color of the indicator. This was the most heterogeneous group, as some participants trusted in automation and others less so. There were equal numbers of men and women, and some people were experienced with automation. This group complied with the IPLA, highlighting its relevance. The other environmental factors were also considered. The last group, *automation enthusiasts*, preferred to keep the automation activated most of the time. Individuals with this profile may be in a mindset whereby they deliberately set out to experience accidental suspension, in order to understand how the system works. This group who based their decision mainly on road markings quality was composed mainly of men. Many respondents expressed trust in driving automation. They were also more experienced with automated systems. Results indicated that the IPLA was taken into account when deciding whether to maintain the automation, but it was not the only factor. Road markings quality was the second most important factor, confirming users' knowledge about how the automation works and its limitations. Automation enthusiasts analysed the risks of the situation and used the IPLA as a source of complementary information. Meyer (2004) talks about experienced operators who use warning signals as information tools to understand their environment. It appears that these respondents were sufficiently experienced for the IPLA to be considered as an information tool, alongside the other sources of information. This raises the question of how to present the information. These profiles should be borne in mind when designing reliability interfaces. The IPLA can come in many forms. Experienced individuals consider it at the same level as other environmental conditions. This information should not be imposed but accessible.

Our results are coherent with those of Kunze et al. (2019), who showed that color is useful for communicating variations in automation reliability. Green and red were used by Wintersberg et al. (2019). Our results show that green and orange are also relevant colors for conveying reliability. However, providing information solely on the instrument cluster may not be suitable for all profiles. Participants in the automation Skeptics and Compliant with IPLA clusters, who primarily relied on this information, risked not considering environmental factors. If reliability information is only displayed on the instrument cluster, individuals may spend too much time looking at it, and not on the exterior environment. *Automation enthusiasts* ascribed just as much important to the environmental factors as they did to the IPLA. For all three profiles, a display suggesting that automation is likely to be suspended might be more appropriate than a display imposing a takeover. It would allow drivers to understand how the automation works and its limitations. Therefore, peripheral awareness displays should to be considered. Kunze et al., (2019) displayed the reliability information in peripheral vision, allowing drivers to keep their attention on a task, whilst having easy access to reliability information. This way of presenting information would be beneficial for all profiles of respondents.

The IIT method allowed us to investigate compliance with reliability information in relation to different environmental factors. Our results showed that the IPLA was taken into account when using the automation. Previous research on reliability information had yielded the same results (Beller et al., 2013; Helldin et al., 2013), confirming that this is a relevant method for investigating human interactions with an interface and, more generally, with new technologies. Moreover, it also allowed us to identify different profiles. When used in other domains, this method has shown that respondents' profiles are decisive in the integration of the information that is presented to them (Hurgobin et al., 2019). This confirms that profiles should be taken into account when developing an interface. The creation of personas, based on clusters, is useful when developing interfaces, as it allows a whole range of human behaviour to be considered.

4.1. Limitation

The main limitation of this study was that the stimuli were written scenarios. The situations were not encountered in real life. Participants may feel that they would react in a particular way in a specific situation, whereas in reality they might act in another way. However, our results show that weather conditions are an important factor in risk perception, in line with previous research (Brooks et al., 2011; Wille & Debus, 2005). Therefore, the projected behaviours highlighted by this method appear to be close to those encountered in real situations. This method has the advantage of allowing a large number of participants to be recruited from a large geographical region. The cluster of *Automation enthusiasts* was the smallest, with just 34 participants. A larger group of respondents corresponding to this profile would allow us to enhance the quality of our results. A second limit is the absence of consideration of cultural differences and the fact that was a convenience sample. Participants were all French speakers, yet the relation with automated driving may be different in other countries. Another limitation is that our assessment of trust in automation was done after the experiment and not before. Some situations may have been perceived of as problematic for the use of automation and made participants aware of the limitations of automated driving. These situations may thus have modulated initial trust in automation. Finally, our representation of the reliability of automation was binary: green or orange. With more shades of reliability, results might have been different.

4.2. Implication & Conclusion

The IPLA was considered and was a determining factor in the use of automation. Personal factors determined how far this display was used to decide whether automation should be manually deactivated before an accidental and surprising suspension. Initial trust in automation was linked to the decision to deactivate the automation. Integrating this system might allow drivers to anticipate accidental suspensions. However, personal traits should also be considered. People who were the most experienced and who trusted automation more based their use of automation rather on environmental circumstances, than on the IPLA. People who had less trust in automation based their judgment mostly on the IPLA, giving little consideration to environmental factors. They may have blindly complied with the IPLA, owing to their lack of experience with automation. Transparent interfaces would allow these individuals to understand and learn what the problematic situations for automation are. Exogenous information, conveyed via a peripheral awareness display, may be suitable for providing information about the limits of automation. These displays allow information to be provided within the driver's peripheral vision, so that the road remains the focus of attention. Peripheral awareness displays, in the same form as those tested by Kunze et al. (2019), have been used in Level-3 vehicles in previous research. Future research should focus on adapting these displays to Level-2 vehicles. To go further and make the interface accessible to all profiles, another possibility would be to use adaptable human-machine systems (Vidulich & Tsang, 2015). Adaptable systems are designed to adapt the interface to the user's ability. With these systems, it should be possible to adapt the interface to the profiles' needs, in terms of mental workload and situational awareness. For the Automation Skeptics, who probably have an important mental workload, these systems should be able to display the information in peripheral vision when high mental workload is detected, thereby freeing up resources for focal vision and consideration of environmental factors. For automation enthusiasts, such systems would only display reliability information if their situational awareness was reduced. Adaptable human-machine systems should be tested in a simulator. To conclude, compliance with reliability information is influenced by several situational and personal factors. Taking these factors into consideration when designing reliability indicators would help make Level-2 automation safer.

Points clés

- Les conducteurs se conforment aux indications de proximité des limites de l'automatisation pour décider de la désactivation des automatisations.
- Les circonstances environnementales sont prises en compte pour décider de la désactivation des systèmes d'automatisation.
- Trois profils de conducteurs peuvent être dérivés des décisions de désactiver l'automatisation : Les sceptiques de l'automatisation, les conformes à l'indicateur de limite et les enthousiastes de l'automatisation.
- Les principaux facteurs qui différencient les profils des répondants sont le sexe, la confiance a priori dans l'automatisation et l'expérience des systèmes automatisés.

Key points

- Drivers comply with indications of proximity to the limits of automation to decide when to deactivate automation.
- Environmental circumstances are taken into account to decide when to deactivate automation.
- Three profiles of drivers can be derived from the decisions to deactivate automation: Automation Skeptics, Compliant with the indicator of the limit, and Automation Enthusiasts
- The main factors that differentiated profiles of respondents were gender, a priori trust in automation, and experience with automated systems.

CHAPTER 6 – INDICATING THE LIMITS OF PARTIALLY AUTOMATED VEHICLES WITH DRIVERS' PERIPHERAL VISION: AN ONLINE STUDY

With the utility of the IPLA highlighted in the previous chapter, the following subject concerns the design such interface. The experimental study described in this chapter tested a design of an IPLA aimed to be perceived in peripheral vision. It was the subject of a chapter in Advances in Human Aspects of Transportation: Proceedings of the AHFE 2021 Virtual Conference on Human Aspects of Transportation. It was reformatted for the purpose of this manuscript. Its formatting consisted in reducing the introduction elements to avoid redundancy.

Monsaingeon, N. (2021). Indicating the Limits of Partially Automated Vehicles with Drivers' Peripheral Vision: An Online Study. In *Advances in Human Aspects of Transportation: Proceedings of the AHFE 2021 Virtual Conference on Human Aspects of Transportation, July 25-29, 2021, USA* (Vol. 270, p. 78-85). Springer Nature.

Résumé

Cette étude visait à évaluer l'efficacité d'un IPLA pour anticiper les suspensions des systèmes automatisés. L'étude a été réalisée en ligne, en présentant des vidéos représentant des situations dans lesquelles l'automatisation se suspendait. Une interface classique a été comparée à un IPLA dédié à la vision périphérique. Les participants décidaient de l'action à réaliser. Les résultats ont révélé que les participants disposant de l'IPLA ont effectué plus d'actions avant la suspension du système et ont exprimé un plus grand confort psychologique que les participants disposant de l'interface classique. Les participants avec IPLA ont effectué moins d'actions appropriées et l'IPLA a été jugé plus exigeant en ressources cognitives. Ces résultats soulignent la pertinence d'intégrer un IPLA dédié à la vision périphérique, mais celui-ci ne devrait pas inciter le conducteur à désactiver l'automatisation lorsque ce n'est pas nécessaire.

Abstract

Automated systems of partially automated vehicles are able to perform the driving task, but can give back the driver all controls in specific conditions. This study aimed to evaluate the effectiveness of an IPLA to anticipate transitions of control. The study was performed online, presenting videos representing situations in which assistance suspended. A classical interface was compared to an IPLA dedicated to peripheral vision. Participants decided which action to perform. The results revealed that the participants who had the IPLA performed more actions before the system suspended and expressed greater psychological comfort than the participants with the classic interface. The participants with the IPLA performed less appropriated actions and the IPLA was rated as more cognitively demanding. These results highlight the pertinence of integrating an IPLA dedicated to peripheral vision, but should not encourage the driver to deactivate assistance when not necessary.

1. Introduction

As observed in the previous chapter (Chapter 5), an IPLA revealed to influence the decision to deactivate assistance in several situations that might be confusing for the driver (e.g., road with bends, bad weather, unclear road markings. Kunze et al. (Kunze et al., 2019) tested reliability indicators in peripheral vision for highly automated vehicles, which allowed the driver to reduce take over time without looking off the road. Interfaces presented in peripheral vision, adapted to partially automated vehicles, would help drivers to perform their supervisory task by reacting appropriately in takeover situations, without looking off the road. This study aimed to evaluate the effectiveness of an IPLA, meant to be displayed in peripheral vision, in deciding on appropriate driving behaviour in hazardous situations of partially automated vehicles.

An IPLA should give information through a gradual display, indicating approach from or withdrawal from limits of assistance. This way, the drivers can gauge their responses depending on the emergency of the situation (Kunze et al., 2019). It has been shown that it is more efficient when an IPLA is state center, informing on the state and intention of action of assistance (Noah et al., 2017). This information needs to be conveyed continuously in order to increase trust in assistance (B. D. Seppelt & Lee, 2007a). Such indicators should also not cause cognitive overload with too much additional information (Mankoff et al., 2003). A way to avoid that is to distribute the information on different sensory channels (e.g., auditory, focal vision and peripheral vision; Politis et al., 2014). Interfaces displaying information in peripheral vision allows the driver to free resources of focal vision to focus it on the road. With peripheral displays, IPLA should include color hue variations and size variations (Kunze et al., 2019). The temperature metaphor appears to be efficient to reflect changes in the urgency of a situation, blue reflecting a passive situation and red reflecting a high degree of urgency (Davis et al., 2017). Distinctive color steps were reported as efficient for the user to better estimate the urgency of the situation (Faltaous et al., 2018). The changes of color, and therefore urgency, should reflect the state of assistance. Finally, information presented at the top of the cluster should be perceived in peripheral vision (Terken et al., 2013).

In this study we attempted to answer the following research question: is the IPLA dedicated to peripheral vision, which follows the requirement described above, more adapted than a classical interface to respond to confusing situations? To answer this question, a group using the IPLA interface was compared with a group using a classical interface, referred here as the Reference interface. The following hypotheses were tested during a video-based online study:

(1) More decisions to perform actions will be taken before the suspension of assistance with the IPLA interface; (2) The selected actions will be more adapted to the situation with an IPLA interface; (3) the IPLA interface will induce a better psychological comfort; (4) the IPLA interface will induce a more important cognitive load.

2. Method

2.1. Participants

The participants were unpaid French-speaking volunteers, recruited via online social networks and mailing invitations. Half of the participants were coworkers from Renault Technocentre or IRT SystemX. The other half were students from the University of Technology of Compiègne, acquaintances, or family members. The sample was composed of 93 participants, including 30 women. Ages ranged from 18 to more than 65 years, the majority of participants aging from 35 to 49 years. The goal of the study was described as the evaluation of a display that aimed to make vehicles' assistance safer. The participants were required to hold a driving licence and to perform the experiment from a computer. The study was explained in the online survey and all participants then completed an informed consent form. Full anonymity was respected.

2.2. Videos

Videos were presented to the participants. The videos depicted four situations from the point of view of a driver of a car in a simulated environment (see Fig. 1). The situations used were reported as situations in which the system can reach its limits in the Renault Clio 5 2019[3] car user manual. They were selected depending on the conditions, assistance could stop functioning. Four situations were depicted: *a road with bends*, *a traffic jam*, *a foggy area* and *an area where road markings were of bad quality*. For each of these situations, two videos were presented: one in which the vehicle's assistance suspended and one in which the vehicle assistance stayed active. This resulted in the presentation of 8 videos (see Appendix C for links to the videos). During the *road with bends* videos, when assistance suspended, the vehicle went

[3] see https://fr.e-guide.renault.com/fra/Clio-5/Assistant-Autoroute-et-Trafic, retrieved on May, 12th, 2020.

off the road in the bend. During the *traffic jam*s video, the vehicle of the participant braked until it reaches its maximum deceleration, then the emergency braking activated. In the *foggy area* videos, assistance suspended due to too much fog density. For the *road markings* videos, road markings were erased, which resulted in the suspension of assistance. In the videos, the tested interface was embedded in the cluster and zoomed so that participants could clearly see the information presented whatever the size of their computer displays (see Figure 16). The videos stopped when the ego vehicle passed the situation or a few seconds after assistance suspended.

Figure 16

Screenshot of a presented video during the experimental task for the Reference group. The road scene is represented, as well as the interface.

2.3. Interfaces

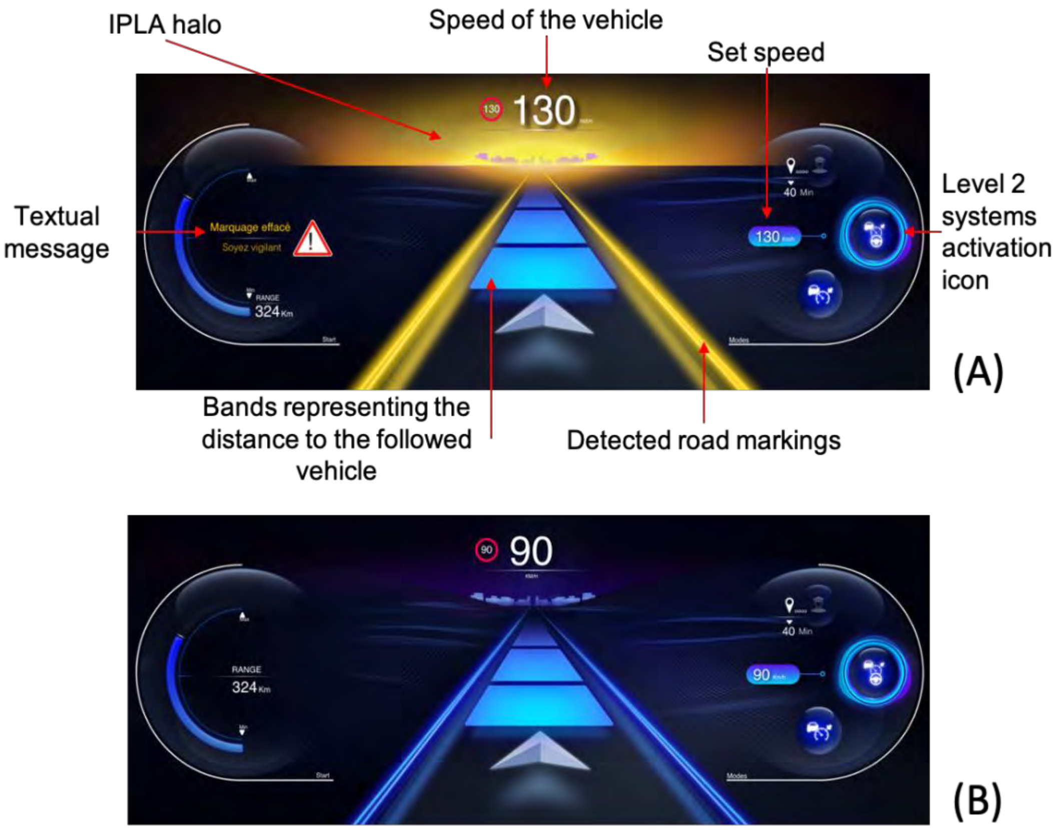
Two interfaces in the instrument's cluster were compared: an interface equipped with an IPLA and a Reference interface (see Figure 17). Both interfaces shared mutual characteristics, as follows. The activated level of assistance was displayed at the right of the screen in blue. When the assistance suspended, the icon turned grey. The detected road markings were displayed in blue on the screen, as well as horizontal bands representing the distance to the vehicle ahead.

Both elements turned grey when assistance suspended. On the left of the cluster, an area was dedicated to textual messages. The Reference interface displayed only the elements cited previously. The IPLA interface featured additional elements. A glowing halo was displayed at the horizon and represented the proximity to the limits of assistance. It could have three representations: blue and narrow, meaning that assistance is functioning as expected; yellow and medium-sized, meaning that assistance is getting close to its limits but won't disconnect yet; red and large, meaning that assistance is close to its limits of proper functioning and suspension is very likely to occur. The halo was located in the upper part of the cluster to be perceived in peripheral vision while looking at the road. In addition to the halo, the central zone of the cluster changed depending on the encountered situation. In the *road with bends* videos, the detected road markings were bent and had the same color as the halo. In the *traffic jam*s videos, the bands representing the distance to the vehicle ahead were the same color as the halo. The *foggy area* videos, the bands and detected road markings were the same color as the halo. In the *bad quality road marking* videos, the detected road markings were the same color as the halo. Finally, the textual zone was used to inform the participant of the cause of the approach to the limits of assistance (e.g., erased road markings) and the appropriate action to be taken.

Figure 17

Screenshots of interfaces presented during the bad quality road markings video.

(A) : IPLA; (B) : Reference interface.



2.4. Task

The participants had to watch videos and to project themselves in the presented situations, as if they were driving. They had to choose between the following decisions of action when they felt the need to do so, or they could wait for the end of the video: brake, turn the steering wheel, decrease set speed, or deactivate the assistance system. If the participants decided to perform an action, the video stopped.

2.5. Procedure

The participants were randomly assigned to either the IPLA group or the Reference group. They were explained that the vehicle was equipped with an assistance system capable of automatically adapting its speed, lane position, and dealing with bends, but the driver was responsible for monitoring the road, and had to keep their hands on the steering wheel. The assistance system was described as having limited capacities, and that the driver had to take back control if limits were reached. In the IPLA group, participants were instructed that to compensate for the assistance system's limitations, the manufacturer had integrated an indicator. The functioning of the indicator was described. An initial familiarization phase then began and featured two videos in which the vehicle passed next to a highway exit. In one of the two videos, the system suspended, in the other one, it did not. The subsequent experimental phase began with the 8 videos, presented in a random order. After each video, the participants responded to questions about psychological comfort. After watching all videos, they rated the amount of visual information displayed on the cluster and completed a sociodemographic questionnaire. The survey was implemented on Qualtrics and took less than 15 minutes to be completed (M = 12.86, SD = 4.23). Full anonymity was respected.

2.6. Measures & Analysis

Several measures were gathered in this experiment. The first measure was the percentage of participants that decided to perform an action before the system was suspended. A second measure regarded the quality of action in the presented situation. Each action of the participants during the videos were classified as "appropriate" or "inappropriate". The coding was decided before the experiment with two experts in assistance systems. An action was classified as inappropriate if it caused an unnecessary suspension of the assistance system or if the action was too late to keep the driver safe. An inappropriate action was classified as "too early" if it caused an unnecessary suspension of the assistance systems. An action was classified as appropriate otherwise. The percentage of appropriate actions was calculated depending on the interface group. A third measure was the psychological comfort experienced after each video, rated by the participants in response to the question "How did you feel during the seconds preceding your action?" on a 6-point scale ranging from 1 (*Not good at all*) to 6 (*Very good*). Finally, the cognitive load caused by the interface was rated by the participant with the following question "Of all the trips you took, how would you rate the amount of information displayed on the cluster?" on a 5-point scale ranging from 1 (*Too small*) to 5 (*Too much*), asked at the end of the experiment. The IPLA and Reference groups were compared using independent student t-tests.

3. Results

The analysis of the percentage of taken action before suspension of the assistance revealed that in the *road with bends* video in which the system suspended, the participants with the IPLA did more actions (M = 75%; SD = 44%) than the participants with the Reference interface (M = 11%; SD = 30%; $t(91) = 8.33$; $p < 0.001$). The analysis on the other videos did not reveal significant differences between the interface groups ($p > .05$).

The analysis of the quality of the performed actions revealed several significant differences. The participants with the Reference interface did more appropriate actions than the participants with the IPLA during *the road with bends* video with the suspension of the assistance ($t(91) = -5.2$; $p < 0.001$), without suspension ($t(91) = -5$; $p < 0.001$), during the *traffic jam*s video without suspension of assistance ($t(91) = -2.2$; $p = 0.03$) and during the video of *bad quality road markings* without suspension of the system ($t(91) = -2.4$; $p = 0.01$). However, the participants with the IPLA did more appropriate actions in the *foggy area* video with the suspension of assistance ($t(91) = 2.4$; $p = 0.01$; see Table 17). The analysis of the other videos did not reveal significant differences ($p > .05$). When looking at the number of "too early" inappropriate actions during the *road with bends* videos with suspension, a greater number of participants with the IPLA interface did "too early" inappropriate actions (n = 26) than the participants with the Reference interface (n = 0. For the *road with bends* video without suspension, more participants with the IPLA interface did "too early" inappropriate actions (n = 17) than the participants with the Reference interface (n = 2).

Table 17

Mean percentage of appropriate action (SD) depending on the interface group and video.
*** p < .05; *** p < .001*

Situation	State of automation	IPLA group	Reference Group
Road with bends	Suspension	27% (45%) ***	75% (43%)
	No suspension	50% (51%) ***	92% (28%)
Traffic jam	Suspension	62% (49%)	63% (49%)
	No suspension	62% (49%) ***	82% (39%)
Foggy area	Suspension	43% (50%) **	20% (41%)
	No suspension	91% (29%)	98% (14%)
Bad quality road markings area	Suspension	61% (49%)	76% (43%)
	No suspension	84% (37%) **	98% (14%)

The analysis of psychological comfort revealed a significant difference between interface groups for the *road with bends* videos when the assistance suspended ($t(89) = 3.1$; $p = 0.002$). The participants with the IPLA experienced a better psychological comfort ($M = 3.88$; $SD = 1.14$) than the participants with the Reference interface ($M = 3.00$; $SD = 1.50$). There was a similar difference for the *road with bends* videos when the assistance did not deactivate ($t(50) = 2.2$; $p = 0.02$). The participants with the IPLA experienced a better psychological comfort ($M = 4.10$; $SD = 1.01$) than the participants with the Reference interface ($M = 3.4$; $SD = 0.68$). The analysis of the other videos did not reveal any significant differences ($p > .05$).

The analysis of cognitive load revealed a significant difference between the interface groups. The participants with the Reference interface ($M = 3.43$; $SD = 0.94$) rated that their interface was less cluttered with information compared to the participants with the IPLA ($M = 4.09$; $SD = 0.98$; $t(91) = 3.3$; $p = 0.001$).

4. Discussion

The participants of the IPLA group did more actions before the suspension of the assistance in the *road with bends* video. This reflects the fact that the IPLA impacted the decision of the participants to perform an action. Hypothesis (1) was therefore verified, but only for this video. This revealed that the IPLA helped the participants to react before the system suspended, potentially avoiding confusion. The halo was perceived and impacted the action of the participants, which follows the results of Kunze et al. (2019) with highly automated vehicles. However, the participants' actions were not always the most appropriated. Hypothesis (2) was partially verified, because the participants of the Reference group did more actions appropriate to the situation compared to the IPLA group for *the road with bends* videos with and without the suspension of the assistance, for the *traffic jam*s video with a suspension, and for the *bad quality road markings* video without suspension. This could be explained by the fact that the halo used only a three-color gradations. Drivers being cautious, they might rather perform an action that will assure them safety, even if it implies deactivating assistance when IPLA indicated a medium approach to the limits. This is confirmed by the fact that a greater number of participants did "too early" inappropriate actions in the IPLA group, causing unnecessary suspension of assistance. Following the recommendations of Kunze et al. (2019), IPLA should be gradual. In our case, the IPLA might have not been gradual enough, presenting alerting information too early relative to the approaching event. However, participants with the IPLA did more appropriate actions during the *foggy area* video with the suspension of assistance. The conveyed information regarding bad weather may have reassured the participants and helped them to react properly. Regarding psychological comfort, Hypothesis (3) was partially verified. During the seconds preceding the participants' decision of action, participants with the IPLA reported better comfort compared to participants with the Reference interface for the *road with bends* video, whether the system suspended or not. Our results follow those of Beller et al. (2013), the IPLA made participants feel more psychologically comfortable, increasing acceptability of the assistance. Interestingly, this is the case for videos in which participants did not react appropriately to the situation. Finally, regarding cognitive load, Hypothesis (4) was verified. The participants with the IPLA evaluated the amount of information of the interface as more substantial than the participants with the Reference interface. However, both groups rated their interface as having too much information displayed (at least 3.4 out of 5). The amount of information given by the IPLA seems to be too important, with the interplay of different parts of the cluster. It should be reduced in order not to cause cognitive overload (Davis et al., 2017). The main limit of this study was that it was performed online, through videos. The reactions of the participants to interfaces and the size of computer displays are not exactly representative of reality. Simulator experiments would allow to evaluate the efficiency of the IPLA in more ecological situations.

Points clés

- La proximité des limites de l'automatisation peut être indiquée en vision périphérique avec des variations de couleur et de taille d'un affichage.
- L'indicateur de proximité des limites de l'automatisme augmente le nombre de décisions de désactivation des systèmes automatisés avant leur suspension.
- Le choix de l'action n'est pas toujours adapté à la situation lorsque l'indicateur de proximité des limites de l'automatisme est affiché.
- La durée présentation des informations de l'indicateur devrait être raccourci et les éléments d'information centralisés.

Key points

- The proximity to the limits of automation can be indicated in peripheral vision using color and size variations.
- The indicator of proximity to the limits of automation increases the number of decision to deactivate automation before their suspension.
- The quality of choice of action is not always better adapted when displayed the indicator proximity to the limits of automation
- The moment of presentation of information from the indicator should be shortened and elements of information centralized.

Chapter 7 – Earcons to reduce mode confusions in partially automated vehicles: Proposition and application of an evaluation method

The goal of this chapter is to evaluate the specific capacity of earcons to influence mode awareness in during transitions of the mode of partially automated vehicles. Earcons of actual partially automated vehicles were related to mode confusions. This chapter's study presents an evaluation method that aims to verify the efficiency of earcons to inform on the mode of automated systems. This method is based on Endsley's model of mode awareness and verify its first two levels: perception and comprehension of the earcons. Earcons indicating the mode of automation thanks to pitch, timbre and number of notes variations were evaluated using the proposed method. This method allowed to ensure the potential of earcons to positively influence mode awareness, which led them to be included in the multimodal interface. The methodology and experiment presented in this chapter is the subject of an article aimed to be submitted to a journal.

Monsaingeon, N., Caroux, L., Langlois, S., & Lemercier, C. (submitted). Earcons to reduce mode confusions in partially automated vehicles: Proposition and application of an evaluation method.

Résumé

Dans les véhicules partiellement automatisés, des confusions de mode peuvent se produire lorsque les conducteurs ne perçoivent pas ou ne comprennent pas le mode d'automatisation indiqué par l'interface. Des earcons reflétant la hiérarchie des modes d'automatisation grâce à la hauteur des sons, à leurs rythmes, et au nombre de variations des notes devraient permettre d'améliorer la conscience des modes. L'objectif de cette étude est de proposer et d'appliquer une méthode d'évaluation de l'efficacité de earcons pour indiquer le mode d'automatisation, en se basant sur un modèle théorique de la conscience de la situation. La méthode d'évaluation consiste à évaluer que les earcons sont perçus correctement (expérience 1), compris de manière isolée (expérience 2), et compris pendant l'exécution d'une tâche visuelle reproduisant les demandes de la conduite automobile (expérience 3). La méthode d'évaluation a été appliquée à des earcons indiquant la hiérarchie des modes d'automatisation. Les résultats de ces expériences suggèrent que les earcons ont été efficacement perçus et compris, en isolation comme pendant une tâche visuelle. Les paramètres des earcons décrits ici peuvent être utilisés pour en produire de nouveaux qui réduiraient les confusions de modes. La méthode d'évaluation des earcons peut être exploitée pour d'autres signaux auditifs afin de s'assurer qu'ils alimentent correctement toutes les dimensions de la conscience des modes.

Abstract

In partially automated vehicles, mode confusions can occur when the drivers do not perceive or comprehend the mode of automation indicated by the interface. Earcons reflecting the hierarchy of the mode of automation using pitch, rhythm and the number of note variations should increase mode awareness. The goal of this study is to propose to apply an evaluation method based on a mode awareness model that aims to assess the efficiency of earcons to indicate the mode of automation. The evaluation method consists in evaluating that the earcons are perceived correctly (Experiment 1), comprehended in isolation (Experiment 2), and comprehended considering the execution of a parallel visual task mimicking driving (Experiment 3). The evaluation method was applied to earcons indicating the hierarchy of modes of automation. The results of these experiments indicated that the earcons were efficiently perceived, comprehended in isolation and during a visual task. The parameters described here can be used to create earcons that would reduce mode confusions. The method of evaluation of earcons can be exploited to other auditory signals to ensure that they correctly feed all dimensions of mode awareness.

1. Introduction

As mentioned in Chapter 4, auditory signals can be used in partially automated vehicles inform drivers about the state of the different automated systems. This avoids them having to gaze at the instruments' cluster, especially if they are already navigating in an infotainment system (Tardieu et al., 2015). *Earcons* (i.e., abstract sounds that are usually arbitrarily linked to the meaning) can convey information regarding the mode of automation to drivers with only minimal training (Kramer, 1994). However, the presence of earcons in partially automated vehicles has been linked to mode confusions (Monsaingeon et al., 2021), and the design of these earcons can impact their recognition (Brewster et al., 1993). In the present study, we explored how earcons reflecting the hierarchy of the automated systems can ensure that automation modes are correctly identified. More specifically, we developed a three-step method to test the efficiency of auditory signals in ensuring mode awareness. The three steps were (1) assessing that the auditory signals were correctly perceived, (2) assessing that they were correctly comprehended in isolation, and (3) assessing that they were correctly comprehended during a parallel visual task mimicking driving.

1.1. Auditory Interface for Level-2 Vehicles

Transmitting information through sound is referred to as *sonification* (Tardieu et al., 2015). Sonification can take the form of either earcons, auditory icons, or spearcons (speech-based earcons; Kramer, 1994). *Earcons* are abstract sounds that are usually arbitrarily linked to the meaning. *Auditory icons* directly use the sound of an object or event to represent its meaning (e.g., noise of tires on rumble strips to represent a lane departure). Unlike earcons, auditory icons usually require no training. *Spearcons* are created using text-to-speech synthesis. Sonification has been used before (Jeon, 2019) in conditionally automated vehicles (SAE Level 3) for takeover requests (TORs). In Level-2 vehicles, more information may be needed about automation mode, such as which system is active and which is not, raising questions about the adequate sonification for Level-2 vehicles. Auditory icons need to directly represent their meaning, but it seems difficult to represent the state of an automated system in this way.

Spearcons are easily understandable, but in the study of Jeon (2019), they were also judged to be more annoying. In Level-2 vehicles, LCA may often have to be suspended because of unclear road markings, leading to frequent signals from the interface. In the present study, we therefore investigated earcons in order to develop auditory signals that would be acceptable to drivers and suitable for Level-2 vehicles. Earcons are already present in some vehicles. However, two studies using similar vehicles with identical earcons both reported mode confusions (Banks et al., 2018; Monsaingeon et al., 2021) Two earcons were used to indicate mode transitions from Level 0 or Level 1 to Level 2, and from Level 2 to Level 1 or Level 0. When drivers had to rely solely on the earcons, they were unable to say whether the vehicle was in Level-0 or Level-2 mode. They might have correctly perceived the earcons, but they were unable to identify the automation mode, hence the mode confusion. Two earcons to represent three possible mode transitions may therefore not be sufficient. We posited that using three distinctive earcons to precisely indicate the new mode following a transition would reduce the risk of mode confusion. Moreover, previous studies have shown that earcons impact reaction times in a visual task (Lemmens et al., 2000). It is therefore important to ensure that earcons do not have a major detrimental effect on the driving task, and are correctly perceived and comprehended.

1.2. Earcons Indicating the Mode of Automation

Earcons can be used to convey several types of information, including how to navigate hierarchical menus. According to the guidelines of Brewster et al. (1993), drivers can easily learn to recognize earcons. Pitch and rhythm are two sound parameters that can be manipulated to help distinguish between earcons and efficiently transmit information. In the context of menu navigation, earcons can be used to represent a hierarchy. Relying solely on earcons, it is possible to locate a position in a hierarchy 97% of the time (Brewster et al., 1998). Sonification has also been shown to help navigate hierarchical menus during a dual task (Jeon et al., 2009). Moreover, it reduces the number of visual fixations on menu displays (Tardieu et al., 2015). Using ADASs in a vehicle with several automation modes can be pictured as navigating in a menu with a hierarchical structure. By pressing activation buttons, drivers can navigate between automation modes, increasing or decreasing the degree of driving assistance. In the present study, we designed earcons to represent the hierarchy of automation modes, by manipulating the parameters of pitch, rhythm, and number of tones. Given previous results for menu navigation, we expected these earcons to allow drivers to recognize the automation mode, even when performing a dual task.

1.3. Research Objective

Earcons have been used in human-machine interactions for over three decades, and over this time they have become increasingly efficient (Brewster et al., 1993). In the automated driving domain, earcons can improve situational awareness, which is usually assessed through questionnaires (Beattie et al., 2015; Gang et al., 2018; Nadri et al., 2021). Even these questionnaires have been standardized and validated, they nonetheless constitute subjective assessments of situational awareness. In the context of automated driving, earcons have to convey information in safety-critical situations that require quick reactions. Their efficiency therefore needs to be objectively measured, to ensure that they trigger the right behaviours. The main goal of the present study was therefore to develop and test a method for objectively verifying that auditory signals have a positive impact on mode awareness.

We viewed mode awareness as a subcategory of situational awareness. We chose the model of situational awareness in dynamic situations of Endsley (1995) as our theoretical framework, rather than the model of cognitive control dynamics Hoc and Amalbertie (2007), which is derived from skill-rule-knowledge model of Rasmussen (1983). The latter concerns the processing of different signals, namely signs and symbols. *Signs* designate concrete or ecological signals, and *symbols* designate abstract signals. As earcons are composed solely of abstract symbols that require learning, only dimensions of this model that are involved in the processing of symbols would be applicable to earcons, whereas all parts of Endsley's tripartite model can be applied to them. As stated by Endsley (1995), situational and mode awareness depend on the perception, comprehension and projection of the automation mode (1995). We tested the ability of the assessment method described here to verify that the earcons representing the hierarchy of automation modes were both correctly perceived and comprehended. We assumed that earcons meeting the guidelines of Brewster et al. (1993) would ensure correct differentiation of signals. We further assumed that earcons that allowed drivers to locate the mode in the hierarchy of modes would ensure correct comprehension of the current automation mode.

We carried out three experiments. Earcon perception was assessed with a same/different task (Experiment 1). Same/different tasks serve to check that different signals are discriminated and differentiated. We expected earcons designed according to the guidelines of Brewster et al. (1993) to be correctly differentiated from each other. We explored earcon comprehension in two steps. First, we verified that the earcons were correctly associated with their meaning. This was evaluated with a cued recall task (Experiment 2). Cued recall tasks serve to evaluate the strength of the relation between two elements in memory. We evaluated how far a cue (i.e., earcon) allowed participants to retrieve its meaning (i.e., automation mode). We predicted that the automation modes associated with the earcons would be retrieved quickly and efficiently. Second, we verified that the earcons were correctly associated with their meanings, relative to the operator's goals. This was tested during a cued recall task performed in parallel with a visual task (Experiment 3). This dual-task paradigm allowed us to assess the degree to which these two tasks affected one another. Participants had to perform a the visual task as well as possible, while retrieving the meaning of the earcons. We predicted that the presentation of the earcons would have a moderate impact on performance on the visual task. We also predicted that meaning of the earcons would be correctly and rapidly recalled during the visual task, but the presence of the visual task would have a moderate impact on recall. Positive results would enable us to conclude that earcons can reduce mode confusions in partially automated vehicles and thus reduce the safety risks for drivers.

2. Experiment 1: Evaluation of the Perception of the Earcons

The first phase of the assessment consisted in evaluating the perceptual level of mode awareness. According to Brewster et al. (1993), earcons that use variations in pitch, rhythm and number of notes can be readily distinguished and efficiently convey information. We therefore tested earcons that met these parameters. A same/different task was performed during which pairs of earcons were played and participants had to decide whether they were identical or different. We expected participants to be able to correctly differentiate between the different earcons.

2.1. Method

2.1.1. Participants

Participants were 35 (M_{age} = 39.94 years, SD = 12.79) unpaid French-speaking volunteers who were recruited online via acquaintances, family members, and social media. Those participants who were recruited via social media belonged either to groups set up to answer questionnaires or to student groups. The goal of the study was described as the assessment of an auditory display designed to make automated vehicles safer. Participants were told that to participate, they needed to hold a valid driver's license. They were informed that they would have to listen to earcons and press keys on a computer keyboard. They were each given the name and contact details of the principal investigator.

2.1.2. Task

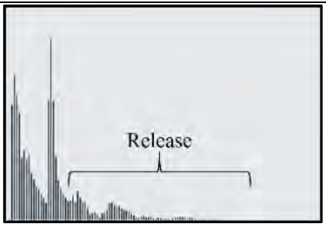
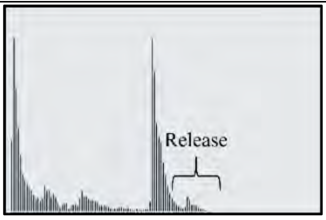
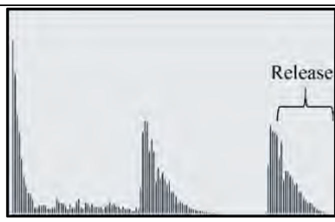
To gauge whether earcons were correctly perceived and differentiated, we administered a same/different task (Farell, 1985). One earcon was presented, and was immediately followed by a second earcon. Participants had to judge whether each pair of earcons was identical or different by pressing a key on their keyboard. An initial training phase consisted of one trial for each type of earcon pair ($N = 9$). Feedback indicated whether the response was correct, and provided the correct answer if a mistake had been made. The subsequent test phase featured five trials for each identical pair ($n = 3$) of earcons, and 10 trials for each different pair ($n = 6$) of earcons, making a total of 75 trials.

2.1.3. Material

The experiment was run on the Qualtrics website, using JavaScript to present the stimuli and allow the participants to respond on their keyboards. The earcons were developed by the IRCAM research institute as part of a joint project with IRT SystemX. All the earcons were based on a C note and lasted around 900 ms. Parameters of rhythm, pitch and number of tones were manipulated, ensuring adequate distinctiveness of the earcons, according to the guidelines of Brewster et al. (1993). In addition to these parameters, attack, decrease, sustain and decay, tone, and duration were manipulated. The earcons were designed to highlight the hierarchy of automation modes, and each one represented the outcome of a mode transition. Pitch, rhythm and release differences were used to illustrate an increase or decrease in the level of automation. An earcon with rising notes, fast rhythm, and long release illustrated an increase in the level of automation. Earcons with descending notes, slower rhythm, and short release illustrated a decrease in the level of automation. The number of tones in the earcon was used to indicate the exact level of automation. Three automation mode transitions were represented by the earcons: transition from Level 0 to Level 2, representing the activation of ACC and LCA; transition from Level 2 to Level 1, representing the suspension of LCA; and transition from Level 2 to Level 0, representing the suspension of both ACC and LCA (see Table 18 for a description of each earcon and Appendix D for a link to hear it).

Table 18

Description of characteristics of earcons.

Earcon label	Automated mode transition	Tones and release description	Notes	Amplitude over time
L2	From Level 0 to Level 2	Two close rising notes. Ascending one octave. Long release.	C2 – C4	
L1	From Level 2 to Level 1	Two spaced out descending notes. Descending one octave. Short release.	C4 – C3	
L0	From Level 2 to Level 0	Three spaced out notes, two descending, and a repetition of the second note. Descending from two octaves with repetition of the last note. Difference in timbre. Very short release.	C4 – C2 – C2	

2.1.4. Measures and Analysis

They participants' answers were recorded. Correct answers were coded 1, and incorrect answers 0. The mean rate of correct responses was calculated for each pair of earcons. Participants could either respond *identical* or *different*. Given that there was one correct response for two possible answers, if the participants correctly perceived the differences or similarities in all the pairs of stimuli, their correct response rate should be above 0.5. The statistical analysis consisted of a one-sample *t* test, comparing the mean correct response rate to a null hypothesis fixed at 0.5.

2.2. Results & Discussion

Results revealed that participants' answers were significantly above chance level for all pairs of earcons (see Table 19).

Table 19

Means rate of correct response and t-test results depending on the pairs of earcons.

Pairs of earcons	Mean rate of correct response (SD)	Results t-test
L2 – L2	0.96 (.17)	$t(34) = 15.70$; $p < 0.001$
L2 – L1	0.99 (.04)	$t(34) = 69.00$; $p < 0.001$
L2 – L0	0.99 (.05)	$t(34) = 61.37$; $p < 0.001$
L1 – L1	0.99 (.05)	$t(34) = 61.37$; $p < 0.001$
L1 – L2	0.96 (.13)	$t(34) = 20.287$; $p < 0.001$
L1 – L0	0.98 (.06)	$t(34) = 48.85$; $p < 0.001$
L0 – L1	0.99 (.05)	$t(34) = 61.37$; $p < 0.001$
L0 – L1	0.99 (.03)	$t(33) = 84.00$; $p < 0.001$
L0 – L0	0.98 (.06)	$t(33) = 50.18$; $p < 0.001$

These results indicated that the earcons were correctly discriminated, confirming that the parameters of pitch, rhythm, and number of notes allow earcons to be efficiently differentiated, in line with the guidelines of Brewster et al. (1993). The earcons therefore satisfied the first level of the mode awareness model by being correctly perceived. In order to build efficient mode awareness, earcons also need to be correctly comprehended, which was assessed in Experiment 2.

3. Experiment 2: Identification of the Earcons

The second phase of the assessment consisted in verifying that the earcons representing the hierarchy of automation modes were comprehended by drivers. To assess this dimension, participants performed a task inspired by cued recall tasks (Carpenter et al., 2006). The goal of this task was to present a pair of stimuli and assess whether one stimulus could cue the recall of the other stimulus. The earcons were first associated with visual icons representing their meanings (i.e., automation mode). They were then emitted alone, and participants had to select the visual icon associated with the target stimulus. We expected the meanings of the earcons to be comprehended by drivers and correctly associated with the visual icons corresponding to the modes.

3.1. Method

3.1.1. Participants

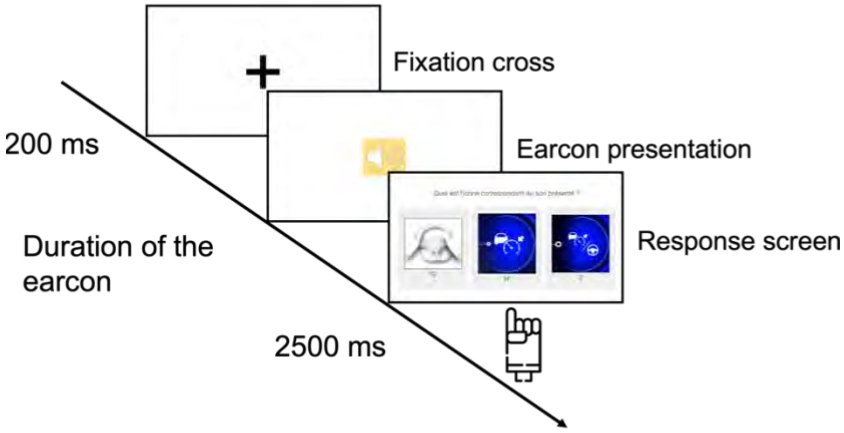
Participants were unpaid French-speaking volunteers. They were recruited by students enrolled on a psychology course at the University of Toulouse Jean-Jaures. Participants could be coworkers, acquaintances, or family members. The goal of the study was described as the assessment of an auditory display aimed at making automated vehicles safer. Participants were told that to take part, they needed to hold a valid driver's license and have experienced cruise control at least once. They were informed that they would have to listen to earcons and press keys on a computer keyboard. The name and contact details of the principal investigator were given to each one. The sample was composed of 528 participants, including 275 women. Ages ranged from 18 to 81 years (M = 37.23, SD = 12.89). Mean length of driving experience was 18 years (SD = 13). Regarding automation, 351 participants had already used cruise control, 82 had used ACC (Level 1), and 36 had already used ACC coupled with LCA (Level 2). Eighty-two participants regularly practiced a musical activity. Some participants reported experiencing tinnitus (n = 20), hyperacusis (n = 7) or another auditory impairment (n = 18), or wearing hearing aids (n = 4). This factor was integrated in the data analysis to control for its effect on performance. More than half of participants were right-handed (n = 333). Of the participants who started the experiment ($N = 1084$), 49% completed the whole survey. All participants filled out an informed consent form. Full anonymity was respected.

3.1.2. Task

The task was a cued recall task (Carpenter et al., 2006). The goal was for participants to correctly associate an earcon with a visual icon representing its meaning. During a study phase, each earcon was associated with a visual icon representing the mode of automation, and participants had to learn the association. In each trial in the test phase, one of three earcons was emitted. A response screen was then presented, displaying the three possible visual icons (see Figure 18). Participants had to choose the visual icon that corresponded to the emitted earcon by pressing the relevant key on their keyboard. The visual icons were always shown in the same order. Once a key had been pressed, the trial ended and a new one began. The participants were instructed to respond as accurately and rapidly as possible.

Figure 18

Example of a trial in the phase with the display duration for each screen.



3.1.3. Material

The experiment was implemented on the Qualtrics website, using JavaScript to present the stimuli and allow participants to respond with their keyboards. The earcons were emitted from either the computer's speakers (n = 314), earphones (n = 96), or headphones (n = 118). The modality in which the earcons were emitted was included in the analysis to take its variance into account. The earcons used in this experiment were identical to those in Experiment 1 (see Table 1 for more explanations). Visual icons were used to represent the three automation modes: Level 2, Level 1, and Level 0 (see Table 20).

Table 20

Visual representation of modes of automation.

Mode of automation	Level-2	Level-1	Level-0
Visual icon			

The task was composed of two phases: a study phase and a test phase. During the study phase, the visual icon of each mode of automation was linked to the corresponding earcon. The duration of presentation was dictated by the participant. Next, each earcon was played, participants had to press the key on the keyboard that corresponded to the correct visual icon. A feedback screen informed them whether they had responded correctly. The earcon-icon association could be shown a second time, if requested by participants. A training phase then began, where a fixation cross displayed for 200 ms was followed by a blank screen during which one of the three earcons was played, then a response screen displaying the three possible visual icons for 2500 ms. Participants pressed a key on the keyboard to choose one of the three visual icons. A feedback screen informed them whether the association was correct, and indicated the correct response if a mistake had been made. This training phase was repeated twice for each earcon, and could be repeated a third time, if requested by participants. The test phase then began. The test trials were identical to the training ones, except that no feedback was given to participants (see Figure 18). There were five trials for each earcon, making a total of 15 trials.

3.1.4. Procedure & Design

Participants began by registering an anonymous number they had been given and signing an informed consent form. The earcons were then played, and participants were invited to adjust the volume of their device to ensure that they could hear clearly. The correct functioning of the computer was tested by pressing each of the keyboard keys that would be necessary for the rest of the experiment. Participants were then given a brief explanation of how Level-2 automated vehicles work, including the automated systems, drivers' behaviours, and visual icons associated with Levels 0, 1 and 2. The study phase came next, followed by the test phase, where the manipulated factor was the earcon that was emitted. It had three modalities (L2 vs. L1 vs. L0) and a within-participants design. At the end of the test phase, a link appeared to Experiment 3, which directly followed Experiment 2.

3.1.5. Measure and Analysis

We measured the quality of responses. This was reflected by the associations that participants made between the earcons and visual icons. Participants' responses were coded 1 when they correctly pressed the key corresponding to the earcon (e.g., *G*, corresponding to the Level-1 visual icon when the L1 earcon was played). Incorrect and absent answers were coded 0. The mean rate of correct answers was then calculated for each condition. The statistical analysis consisted of a one-sample *t* test, and was carried out using jamovi software version 2.0.0.0. Given that there was one correct response out of three possible answers, if participants correctly identified the earcons, their correct response rate would be above 0.3. The normality of residuals of the quality of response was not respected for each condition according to the Shapiro-Wilk normality test ($p > .001$). The nonparametric Wilcoxon's rank-sum test was applied with a null hypothesis set at below 0.3.

3.2. Results

The quality of response was reflected by the median rate of correct answers, depending on the earcon. The median rate of correct answers was 0.80 (0.40) for the L1 and L2 earcons, and 1.00 (0.20) for L0 (see Table 21). For all three conditions, the correct reponse rate was significantly higher than the null hypothesis ($p < .001$).

Table 21

Means (SD) and medians (IQR) of correct response rate depending on the earcon, with the associated Wilcoxon rank-sum tests.

	Mean (SD)	Median (IQR)	Wilcoxon W	*p*
L1 earcon	0.75 (0.29)	0.80 (0.40)	136612	< .001
L2 earcon	0.73 (0.31)	0.80 (0.40)	134941	< .001
L0 earcon	0.86 (0.24)	1.00 (0.20)	139151	< .001

Note. H_a $\mu > 0.3$

3.3. Discussion

Results revealed that participants were able to correctly retrieve the meaning of the earcons, with a success rate of around 75%. These results are consistent with the literature on auditory TORs for Level-3 vehicles. Earcons signaling a change of mode of automated systems can be correctly understood (Jeon, 2019). Earcons representing the hierarchy of automation modes appeared to permit efficient recall of the relevant mode. The comprehension level of Endsley's model of situational/mode awareness (1995) was therefore attained by the earcons we tested. These earcons were designed for Level-2 vehicles, meaning that drivers would hear them while engaging in the driving task. The drivers' goal would therefore be to continue driving safely and understand the auditory information they were given without being distracted. Accordingly, the final step in the assessment consisted in verifying that the earcons were comprehended by drivers pursuing a goal (i.e., performance of a visual task mimicking the driving task).

4. Experiment 3: Identification of Earcons During a Visual Task

The third phase of the study consisted in verifying that the earcons were comprehended and taken into consideration according to the drivers' goals. On the road, a drivers' goal is to perform the driving task safely. A cued recall task similar to that used in Experiment 2 was therefore administered to participants during a visual task that mimicked the attentional demands of a driving task. The aim of this cued recall task was to ensure that the meaning of the earcons was retrieved while participants pursued a goal close to driving. We assessed the impact of earcon comprehension on performances on the visual task. We also measured the quality and speed of recall of the earcons' meaning with and without the visual task, in order to gauge the extent to which the presence of the visual task affected recall. We expected the presentation of the earcons to impact performance on the visual task, but not to reduce it substantially. We expected participants to retrieve the meaning of the earcons during the visual task. Although we predicted that this recall would be poorer when the visual task was present rather than absent, we did not expect this decrease to be critical.

4.1. Method

4.1.1. Participants

The participants of this experiment were the same as those of Experiment 2.

4.1.2. Task

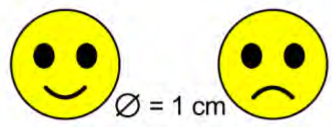
This experiment used a dual-task paradigm inspired by Tardieu et al. (2015). It was composed of a visual detection task and a cued recall task identical to the one in Experiment 2.

Visual Task

The visual detection task was inspired by Tardieu et al. (2015), and aimed to mimic the sustained visual attention required when driving and the need to react to sudden events. Participants had to press a key on their keyboard as fast as possible whenever a target appeared in the visual scene. Each stimulus was displayed for 2500 ms. One stimulus out of four was a target, and the three others were distractors. The order of presentation was random and different for each participant. The background of the screen was white, and the diameter of each stimulus was 1 cm (see Figure 19).

Figure 19

Target (left) and distractor (right) used in the primary task.



Auditive Task

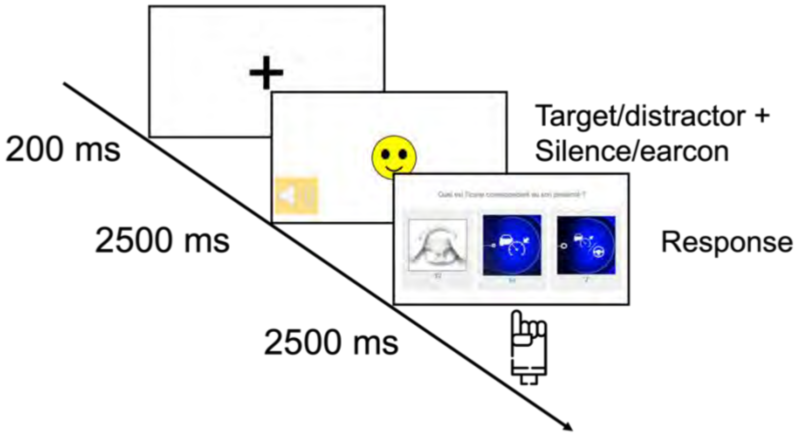
The secondary task was identical to the cued recall task in Experiment 2. During the presentation of the visual task stimuli, either one or none of the three earcons was played at the same time. A response screen was then displayed for 2500 ms. Participants either pressed the key of the visual icon corresponding to the earcon or waited for the response screen to disappear. There was one earcon for every three silences, with a randomized order of presentation.

4.1.3. Procedure

Experiment 3 directly followed Experiment 2, and shared several common features. After clicking on the link at the end of Experiment 2, participants started the training phase of the visual detection task. This involved two blocks of four trials, with the occurrence of one target for every three distractors. Feedback informed participants of the quality of their answer (i.e., green tick for correct answer, red cross for incorrect answer), and indicated the correct response in the case of a mistake. A training phase preceded the dual task in each condition. The instructions specified that the visual detection task had to be performed as accurately as possible, and that similar instructions to those in the previous experiment applied to earcon recall. The correct answer was indicated if a mistake was made. The experimental phase then began (see Figure 20). There was one visual target for every three distractors. It was repeated once for each earcon condition. There was one silence for every earcon. After the experimental phase, participants completed a sociodemographic questionnaire and were thanked for their participation.

Figure 20

Example of a trial in Experiment 3, with the duration of each screen.



4.1.4. Experimental Design

A 2 (within-participants) × 4 (within-participants) experimental design was used. The first factor was the stimuli of the primary task (target vs. distractor). The second factor was the earcon presentation (silence vs. L2 earcon vs. L1 earcon vs. L0 earcon).

4.1.5. Measures & Analysis

In the visual and the auditive task, two measures were gathered: the quality of response and reaction time.

Measure and Analysis of the Visual Task

Regarding the quality of responses, each correct detection of the target was coded 1, and each absent detection was coded 0. The mean correct detection rate was then calculated. Regarding reaction times, participants could start responding 500 ms after the stimulus appeared on the screen, to compensate for differences in earcon duration. We calculated the interval between the onset of each stimulus and each participant's response. All responses within 200 ms, assumed to be involuntary button presses, or after 2500 ms, regarded as errors or Internet connection issues, were excluded. Reaction times were transformed following a logarithmic function, in order to correct the positive skewedness that is usually observed with reaction times (Howell, 2012). Linear mixed models were calculated for both quality of detection and reaction time, to gauge the effect of the earcons. The earcon variable was treated as a fixed factor. To control for factors that potentially impacted performances, participants' handedness, the modality in which the earcon was emitted, and previous experience with ADASs were also included as fixed factors. Participant was a random factor. Bonferroni's post hoc tests were carried out to compare each condition of the earcon factor.

Measure and Analysis of the Auditive Cued Recall Task

We recorded the quality of responses and reaction times for this task. Quality of response was identical to Experiment 2. The mean rate of correct answers was calculated for each condition. Reaction time corresponded to the interval between the completed presentation of the response screen and the submission of the page to the Qualtrics server. This interval represented the time needed by the participant to process the information and respond, as well as the time needed to send the signal to the Qualtrics server indicating that the page had been sent. Participants had 2500 ms to answer, after which the page was automatically sent. The data exclusion criteria were similar to those in Experiment 2. Reaction times were transformed using a logarithmic function, in order to correct the positive skewedness that is usually observed with reaction times (Howell, 2012). We compared participants' performances on the recall of the earcons' meanings between Experiments 2 and 3. This comparison served to estimate the effect of multitasking on 1) the recall of the earcons, and 2) reaction times. The task paradigm (single task vs. dual task) and the earcon factor (L2 vs. L1 vs. L0) were included as fixed factors in a mixed linear model. Fixed factors similar to those in Experiment 2 were also included in the model. Participant was a random factor. Bonferroni's post hoc tests were carried out to compare the conditions of each variable and their interactions.

4.2. Results

4.2.1. Visual Detection Task

The mixed linear model revealed a significant effect of earcon condition on quality of response, $F(3, 1566) = 9.08$, $p < 0.001$. Detection was significantly better when no earcon was present ($M = 0.94$, $SD = 0.15$) than when either the L1 ($M = 0.88$, $SD = 0.26$), L2 ($M = 0.90$, $SD = 0.24$), or L0 ($M = 0.89$, $SD = 0.25$) earcon was present (see Table 22 for detailed tests). No other comparisons were significant.

Table 22

Bonferroni's post-hoc tests of quality of detection between each experimental condition.

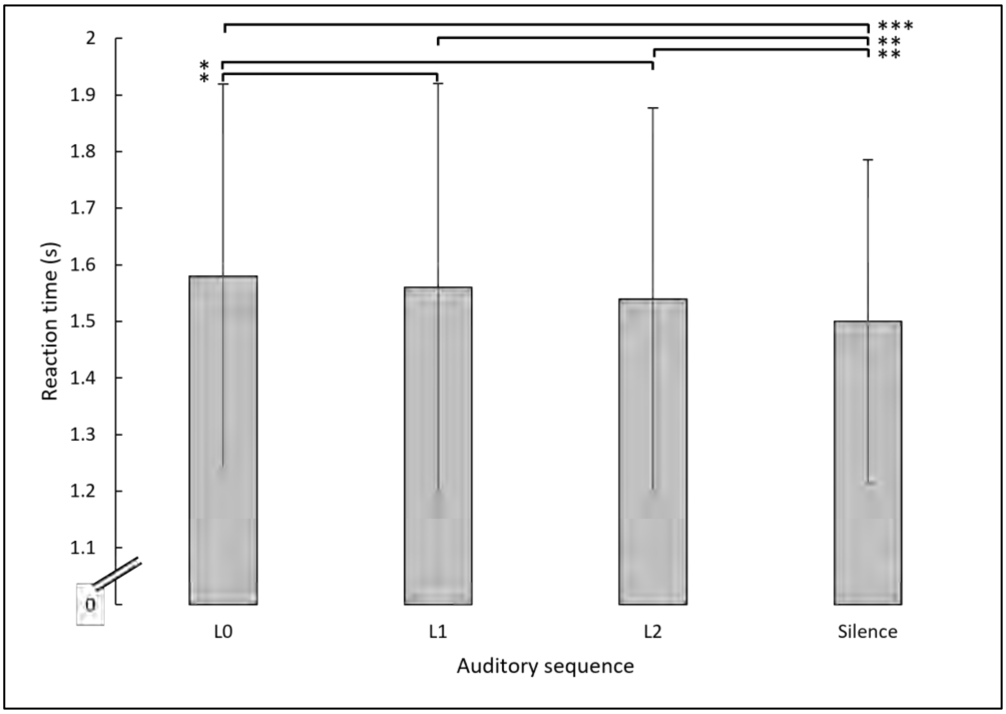
Compared conditions	*df*	*t*	*p*
Silence - L1 Icon	1566	4.96	< .001
Silence - L2 Icon	1566	3.35	0.005
Silence - L0 Icon	1566	3.80	< .001
L1 Icon - L2 Icon	1566	-1.61	0.639
L1 Icon - L0 Icon	1566	-1.17	1.000
L2 Icon - L0 Icon	1566	0.45	1.000

Regarding reaction times, results revealed a significant effect of earcon, $F(3, 1506) = 8.15$, $p < 0.001$ (see Figure 21). According to Bonferroni's post hoc tests, reaction times were significantly faster when no earcon was present ($M = 1.51$, $SD = 0.29$) than when either the L1 ($M = 1.55$, $SD = 0.38$), $t(1500) = -3.74$; $p = 0.001$, L2 ($M = 1.54$, $SD = 0.34$), $t(1499) = -3.37$; $p = 0.005$, or L0 ($M = 1.58$, $SD = 0.34$), $t(1500) = -6.48$; $p < .001$, earcon was present. L0 also differed significantly from both L1 ($M = 1.56$, $SD = 0.36$), $t(1495) = -2.71$; $p = 0.041$, and L2, $t(1495) = -3.10$; $p = 0.012$. L1 and L2 did not differ significantly ($p > 0.1$).

Figure 21

Reaction time to the visual detection task depending on the earcon condition.

** Symbolize a p < 0.5, ** symbolize a p < .01, *** symbolize a p < .001*



4.2.2. Auditive Task: Single Task Vs. Dual Task

We compared performances on earcon recall in Experiment 2 (single task) and Experiment 3 (dual task). The linear mixed model revealed a significant effect of task paradigm on mean correct recall, $F(1, 2612) = 35.70$, $p < 0.001$. Performances were better when the recall was performed in the single-task condition ($M = 0.78$, $SD = 0.29$) than in the dual-task one ($M = 0.73$, $SD = 0.37$; see Table 23 for means in each condition). The effect of earcon was also significant, $F(1, 2612) = 58.079$, $p < 0.001$. No interaction was found between the two.

Table 23

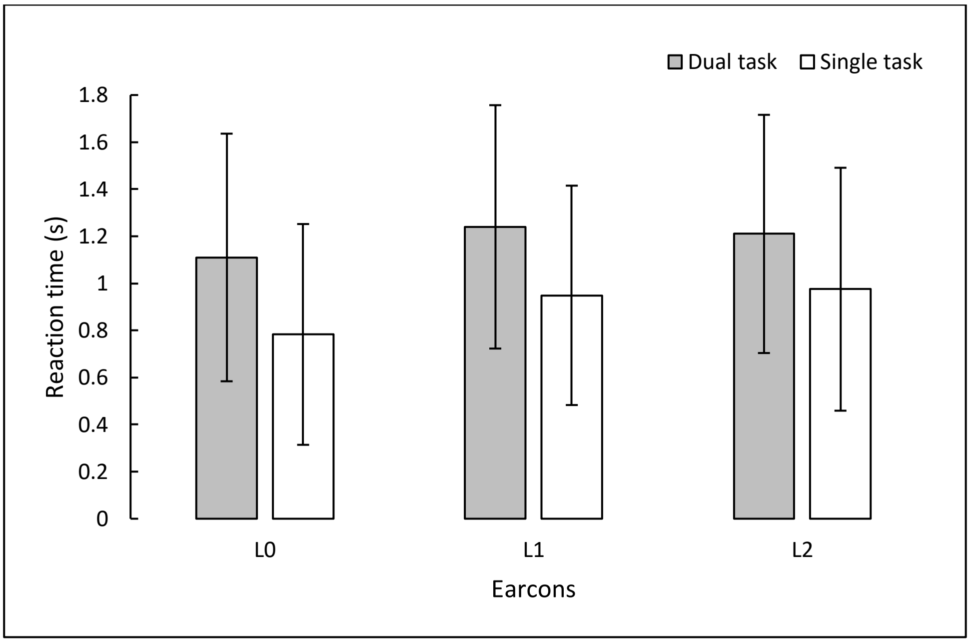
Means of correct recall, depending on the task paradigm and the earcon.

	Earcon		
Task paradigm	L2	L1	L0
Single	0.73 (0.32)	0.75 (0.30)	0.86 (0.59)
Dual	0.70 (0.38)	0.69 (0.37)	0.79 (0.64)

Regarding reaction times, the linear mixed model revealed a significant main effect of task paradigm, $F(1, 2154) = 500.38$, $p < 0.001$. Reaction times were significantly shorter in the single-task condition ($M = 0.90$, $SD = 0.49$) than in the dual-task one ($M = 1.18$, $SD = 0.52$). The main effect of earcon was also significant, $F(2, 2118) = 68.01$, $p < 0.001$. The interaction between the two was also significant, $F(2, 2116) = 5.57$, $p = 0.004$ (see Figure 22). The differences in reaction time according to task paradigm appeared to be greater for L0 (difference = -0.18) than for either L1 (difference = -0.14) or L2 (difference = -0.13).

Figure 22

Reaction time in the auditory recall task, depending on the task paradigm and the earcon.



4.3. Discussion

Results revealed that the earcons reflecting the hierarchy of automation modes were correctly associated with these modes during a visual task. Performances on the visual detection task were impacted by the presence of earcons. There were more correct responses and reaction times were shorter when no earcons were present. These results are coherent with the findings of Lemmens et al. (2000). Although the presence of earcons impacted the visual task, the mean correct response rate only fell by around 5%, and mean reaction times only increased by around 50 ms. Regarding the influence of the dual-task paradigm on the recall of earcon meanings, the mean correct response rate and reaction time were better in the single-task condition (Experiment 2) than in the dual-task one (Experiment 3). The presence of the visual task induced a 5% reduction in the correct response rate and an increase of around 280 ms in the mean reaction time. These results are line with the multiple resources model of Wicken (2008). Although the visual and auditory tasks did not share sensory resources, they did share processing and motor resources, which had a negative impact on performances.

5. General Discussion

The general goal of this study was to develop and test an assessment method to verify that auditory signals positively impact mode awareness. Earcons respecting the guidelines of Brewster et al. (1993) were designed to reflect the hierarchy of automation modes and thereby increase mode awareness. Their potential ability to impact mode awareness was assessed with a 3-step method based on Endsley's (1995) model of situational/mode awareness. Experiment 1, which tested earcon perception, showed that the parameters of pitch, rhythm, and number of notes allowed for efficient differentiation of the earcons. The results of Experiment 2 showed that participants comprehended the earcons in isolation. The results of Experiment 3 showed that the meanings of the earcons were correctly recalled during a parallel visual task that mimicked driving, although this dual-task condition negatively affected both performances on the visual task and recall of the earcons' meanings.

In vehicles currently on the market, mode transitions between manual driving (Level 0), ACC (Level 1), and ACC with LCA (Level 2) are usually signaled by visual icons. The use of a single modality can lead to mode confusion (Monsaingeon et al., 2021). However, some vehicles provide both visual icons and auditory signals to indicate mode changes. The current use of earcons in Level-2 vehicles can also cause mode confusion (Banks et al., 2018; Monsaingeon et al., 2021). This may be because the same earcon is used to indicate mode transitions from Level 2 to either Level 1 or Level 0. More specifically, according to Endsley's model of mode awareness (1995), it may be that these earcons are either not perceived or not comprehended, interfere with the drivers' goal, or do not allow them to project the mode resulting from the transition.

The earcons used in the present study indicated exactly which mode resulted from the transition. According to the guidelines of Brewster et al. (1993), the parameters of pitch, rhythm, and number of notes can be manipulated to create distinctive earcons. Our results confirmed this, as participants successfully differentiated between the earcons (Experiment 1). By varying these earcon parameters to indicate different automation modes, it is therefore possible to achieve the perceptual level of information processing identified in the model of situational awareness of Endsley (1995). Participants were able to rely on variations in the earcon parameters to locate a position in the hierarchy (Brewster et al., 1998). Variations in pitch and rhythm were used to indicate an increase or decrease in automation, while the number of notes allowed participants to identify the exact automation mode. Our results further indicated that using sound parameters to represent the hierarchy of automation modes efficiently allowed participants to recall the earcons' meaning (Experiments 2 and 3), indicating that the comprehension level the situational awareness model (Endsley, 1995) was achieved. To ensure accurate mode awareness, these earcons should also ensure that the third level of the mode awareness model (i.e., projection of future state of automation) is achieved. These earcons accurately represented the mode of automation resulting from the transition. Thanks to them, drivers should therefore be able to project the mode of automation resulting from each transition without having to look at the instrument's cluster. Future studies should investigate the ability of drivers to project the future mode of automation by relying on the earcons.

Even though earcons potentially reduce mode confusion, it is important that they have as little impact on driving performances as possible. The results of Experiment 3 suggest that the earcons had a slightly detrimental effect on the visual task. This raises the question of whether the benefits of having better mode identification outweigh this disadvantage. Performances on the visual task declined by around 5% when the earcons were present, and there was an increase of around 50 ms in the mean reaction time. Monsaingeon et al. (2021) found that three out of 10 participants exhibited mode confusion in a Level-2 vehicle where exclusively visual information was used to indicate automated systems' activation state. The cost-benefit ratio concerning the presence versus absence of earcons in the visual task therefore seems to be in favor of using earcons.

The cognitive processing of auditory signals may not have had a detrimental effect on the main task because of inattentional blindness. In aviation, researchers have found that the latter can occur with auditory signals that are salient and relevant (Dehais et al., 2014). In the study of Dehais et al. (2014), pilots did not perceive an auditory alert signaling that an emergency maneuver was necessary, even though they were familiar with the auditory signals. Cognitive load and task difficulty were assumed to be the cause of their inattentional blindness. However, it was not possible to determine which levels of situational awareness were or were not achieved. Our method makes it possible to examine how auditory signals affect each level of information processing to create situational/mode awareness, the ultimate goal being to refine the signals so that every level is met.

5.1. Limitations

The present study had two limitations. The main limitation was that it was conducted online. We measured reaction times based on an Internet page being sent to online servers. The quality of the Internet connection may therefore have affected this measure. To reduce this bias, the experiment was performed on a large number of participants, and those with extreme reaction times were excluded. A second limitation of this study was the small number of times the experimental conditions were repeated, especially the dual task, which was only repeated twice. Given that participants performed Experiments 2 and 3 one after the other, it was important to minimize the duration of the experiment, in order to avoid dropout. We managed to keep the study duration to around 20 minutes, but still observed a 50% dropout rate. To increase the number of times the experimental conditions are repeated, Experiment 3 should be performed separately from Experiment 2.

5.2. Conclusion & Perspective

We based our assessment of the ability of earcons to inform drivers about ADASs on Endsley's hierarchical model of situational/mode awareness (1995). Perception of the earcons was verified in Experiment 1. Experiments 2 and 3 revealed that the earcons were correctly comprehended, even when participants performed a visual task. The earcons we tested were therefore correctly perceived and comprehended. Further research is needed to validate the third level of the model (i.e., ability to anticipate the future state of the ADASs) in a simulator or on the road. In our experiment, earcons were learned through repeated presentations. However, when individuals buy a new vehicle, there is no such opportunity to learn the earcons. Future studies should examine the possibility of manipulating the earcons' parameters so that they can be learned through usage, without any explanations of their meaning. Once this step is validated, the earcons can be embedded in an interface presenting visual information, in order to assess their acceptability and utility in realistic situations. The use of earcons represents a real opportunity to reduce mode confusion rather than causing it.

Points clés

- Les earcons sont des signaux auditifs abstraits dont la signification doit être apprise.

- Dans des études antérieures, les earcons dans les véhicules partiellement automatisés ont été associées à des confusions de mode.

- Une méthode d'évaluation des earcons pour améliorer la conscience des modes est développée sur la base d'un modèle de la conscience de la situation et repose sur trois expériences

- Les earcons représentant la hiérarchie des modes d'automatisation grâce à la hauteur, au timbre et au nombre de notes s'avèrent stimuler efficacement la conscience des modes.

Key points

- Earcons are abstract auditory signals with a meaning that must be learnt

- Earcons in partially automated vehicles have been related to mode confusions in previous studies.

- A method of evaluation of earcons to improve mode awareness is developed based on a model of situational awareness and is based on three experiments

- Earcons representing hierarchy of modes of automation thanks to pitch, timber, and number of notes reveal to be efficiently stimulating mode awareness.

Chapter 8 – Influence of Haptic Feedback on Mode Awareness in Partially Automated Vehicles

This Chapter presents a two-part study that aimed to evaluate the effect of haptic interfaces on mode awareness. The steering wheel is exploited in 2 experiments because of its capacity to support information regarding the direction of the vehicle. Two feedback in the steering wheel was investigated: tactile feedback (perception on the skin, usually vibration), and kinesthetic feedback (proprioceptive perception of a muscular effort). A similar evaluation method as the one proposed in Chapter 7 is applied in this study to evaluate the capacity of haptic feedback to induce accurate mode awareness by being perceived and comprehended correctly. The results of the study presented in this chapter allow to affirm that the proposed tactile haptic feedback increase detection of suspensions of LCA and that important kinesthetic feedback improve the rapidity of detections. These positive results led to the integration of the haptic feedback of the steering wheel in a multimodal interface evaluated in Chapter 7. The experimental study presented in this chapter was the subject of a research article submitted to the journal Human Factor. The article was reformatted for the purpose of this manuscript.

Monsaingeon, N., Caroux, L., Langlois, S., Wang, J., & Lemercier, C. (submitted). Did You Feel It? Influence of Haptic Feedback on Mode Awareness in Partially Automated Vehicles.

Résumé

Deux types de retours haptiques peuvent être transmis par le volant : les signaux kinesthésiques et les signaux tactiles. Les signaux kinesthésiques font référence à la perception d'un effort musculaire. Les signaux tactiles font référence à la perception tactile de la peau. Les deux types d'informations semblent être des moyens prometteurs pour transmettre des informations rapidement et sans formation. La présente étude vise à évaluer l'effet d'interfaces utilisant des signaux kinesthésiques et tactiles dans le volant sur la détection de suspension des automatisations dans des véhicules partiellement automatisés. Dans la première de deux expériences sur simulateur de conduite, nous avons utilisé une tâche identique/différent pour nous assurer que les conducteurs pouvaient distinguer différents niveaux de signaux kinesthésiques. Dans la deuxième expérience, l'effet de signaux kinesthésiques et tactiles sur la conscience des modes a été évalué sur deux types de routes : une route à une voie comportant des virages et une autoroute droite à deux voies. Les participants avaient pour tâche de détecter la suspension des systèmes automatisés. La qualité de la détection a été évaluée à l'aide d'indices de détection du signal, et les temps de réaction ont été mesurés. Les signaux kinesthésiques et tactiles ont permis d'augmenter la conscience du mode. Les signaux kinesthésiques ont induit une détection plus rapide des transitions de mode. Les signaux tactiles ont réduit les temps de détection sur les routes avec virages par rapport aux routes droites, mais cet effet a été atténué par la présence des signaux kinesthésiques. Nos résultats soulignent la pertinence d'une interface haptique dans le volant pour informer les conducteurs sur les suspensions des automations. Les interfaces haptiques devraient être intégrées dans les véhicules partiellement automatisés afin d'accroître la conscience des modes.

Abstract

Two types of haptic feedback can be provided via the steering wheel: kinesthetic and tactile. Kinesthetic refers to the perception of a muscular effort. Tactile refers to the tactile perception of the skin. Both types appear to be promising means of conveying information quickly and without training. The present study was conducted to assess the effects of this feedback on the detection of mode transitions in partially automated vehicles. In the first of two experiments, we administered a same/different task to ensure that drivers in a simulator could distinguish between different levels of kinesthetic feedback. In the second experiment, the effects of kinesthetic and tactile feedback on mode awareness were assessed on two types of road: single-lane road with bends, and straight divided highway. Participants again drove in a simulator, and had to detect the suspension of automated systems. The quality of detection was assessed using signal detection indices, and reaction times were measured. Both kinesthetic and tactile signals increased mode awareness. The kinesthetic feedback induced faster detection of mode transitions. Tactile feedback reduced detection times on roads with bends versus straight roads, but this effect was mitigated by the presence of kinesthetic feedback. Our results highlight the relevance of haptic feedback for informing drivers about mode transitions. Haptic interfaces should be embedded in partially automated vehicles to increase mode awareness.

1. Introduction

When using automated driving systems, drivers may remain passive for some time, leading to the so-called out-of-the-loop phenomenon (Endsley, 1995). However, the systems may then require these drivers to fully or partially control of the vehicle, depending on the situation. The problem is that drivers may miss this transition if they are out of the loop, as this can lead to reduced awareness of the current state of the system and the situation. Nevertheless, just as a rider is bound to a horse by the saddle and reins (see Flemisch et al., 2003, for the H-Metaphor), so a driver is bound to a vehicle by several physical elements. One of these is the steering wheel. Haptic feedback transmitted through the steering wheel can be perceived quickly and efficiently (see Gaffary & Lécuyer, 2018, for a review). In the present study, we therefore investigated the effect of haptic feedback in the steering wheel to communicate automated systems' current state of the automated systems and any mode transitions. We assessed the effects of two types of haptic feedback: kinesthetic feedback (i.e., perceived exertion of muscular force; Experiment 1) and a combination of kinesthetic and tactile (i.e., perception of a tactile sensation such as a vibration) feedback (Experiment 2). Participants drove in a simulator and were informed of the automation state by tactile and kinesthetic feedback. Our objective was to verify that these haptic interfaces can efficiently inform drivers about the state of automated systems.

1.1. A Different Approach to Driving: Partial Automation

Automated driving may change the way car drivers usually drive. Level-2 automated vehicles have a particular state, because the drivers directly collaborate with the automated systems through the steering wheel. Both the LCA and the driver act on the direction with differing degrees of freedom, depending on the vehicle (i.e., some vehicles are more cooperative than others; see Monsaingeon et al., 2021). The efficiency of this collaboration is crucial for the comfort and safety of drivers and passengers alike. When the situation allows it, drivers can leave the LCA to do most of the work and thus enjoy a more comfortable ride. However, they have to be able to resume full lateral control if the LCA can no longer handle the situation, as is the case when approaching a bend at a high speed (Endsley, 2017) or when there are no clear road markings. In such situations, the LCA relinquishes control with no forewarning. If the driver does not anticipate this type of situation, it can be uncomfortable, if not dangerous, especially if there is a loss of mode awareness, namely, an inability to perceive, comprehend and make projections about automation state (Endsley, 1995). It is therefore vital for the vehicle to quickly inform the driver when the LCA's state changes and a takeover is needed. This role is performed by the car's interface.

1.2. Communication Between Automated Systems and the Driver

Visual displays, using icons on the instrument's cluster, can give drivers two important items of information about automated systems: their current state, and transitions from one mode to another. Icons are displayed continuously on the instrument's cluster when automated systems are active, to inform the driver about the current state. When an automated systems (e.g., LCA) transition from an active mode to an inactive one, the relevant icon usually changes color (see Monsaingeon et al., 2021, for examples). The driver gauges the state of the LCA by identifying the meaning of this icon change (i.e., signal) and ignoring any irrelevant and distracting icon changes (i.e., noise). The quality of drivers' behaviours in critical situations is directly impacted by their ability to distinguish signals from noise. SDT provides a theoretical framework for measuring this ability (Stanislaw & Todorov, 1999): the numbers of times a signal is correctly or incorrectly detected, depending on its presence or absence, are merged into indices. As stated by Janssen et al. (2019), SDT can be applied to evaluate mode awareness in an automated driving context. The degree to which automated systems' state can be distinguished, depending on the interface, can be measured and compared using SDT indices. Interfaces should be designed to induce efficient signal detection concerning automated systems' state, in order to induce adequate mode awareness. The problem is that a instrument's cluster contains many visual icons. Information about automated systems' state therefore risks being buried, making it more difficult for drivers to distinguish between signal and noise. This is especially problematic, given that most available visual attentional resources are already allocated to the driving activity (Sivak, 1996). As stressed by Wickens (2008), one alternative consists in using multimodal interfaces. Distributing the information across several sensory channels increases the chances that drivers will perceive and comprehend the message. The multimodal interface of a Level-2 vehicle that has so far received the most attention is the one used in Tesla's Model S. Visual icons continuously indicate the state of the LCA, and auditory signals indicate a transition of control toward the driver or the vehicle. However, several studies have shown that these signals can be misunderstood (Banks et al., 2018; Endsley, 2017). Other sensory modalities (e.g., haptic) could therefore be used to inform drivers quickly and efficiently.

1.3. Haptic Feedback

Haptic feedback consists in using tactile and/or kinesthetic sensations to convey information (Gaffary & Lécuyer, 2018). This type of feedback has been used in driving in the past, and has the advantage of being quickly perceived, even in very cognitively demanding situations (Murata & Kuroda, 2015; Scott & Grey, 2008). Haptic feedback can be understood more easily than auditory signals, as it resonates with drivers' mental models (Suzuki & Jansson, 2003). Moreover, multimodal interfaces using haptic and visual warning signals tend to reduce reaction times in emergency situations, compared with unimodal interfaces (Politis et al., 2014). Another major advantage of haptic feedback is the possibility of using a wide variety of media, providing physical contact is maintained with the driver. These media include the steering wheel, seatbelt, pedals, seat, dashboard, and clothes (Gaffary & Lécuyer, 2018). The choice of medium may depend on the nature of the message that needs to be conveyed. For example, the steering wheel seems an appropriate means of giving information or warnings related to the car's trajectory, as the signal is emitted from where the action must take place. Warnings in the form of a vibration or a torque in the steering wheel have been shown to be more efficient than a visual equivalent, when it comes to preventing lane departures (Katzourakis et al., 2014; Onimaru & Kitazaki, 2010). Katzourakis et al. (2014) concluded that haptic feedback was not efficient for the low level of automation they tested, but might work better with higher levels of automation. The LCA acts directly on the steering wheel, and therefore corresponds to a higher level of automation. The steering wheel is an effective organ of the vehicle, and when the LCA is activated, the driver and the system both act on it, in collaboration. The steering wheel therefore seems a suitable medium for conveying information about the state of the LCA.

1.4. Haptic Feedback Through the Steering Wheel

We can distinguish between two types of haptic feedback in the steering wheel: tactile and kinesthetic. *Tactile feedback* refers to a tactile perception through the skin (e.g., vibration). It can be used as a brief signal. *Kinesthetic feedback* refers to kinesthetic perceptions of muscular effort (e.g., when a torque is applied in the steering wheel), and can be used to continuously supply information. Both types of feedback have been used to inform drivers about lane departures (Katzourakis et al., 2014; Suzuki & Jansson, 2003). Some vehicles are now equipped with a Lane Departure Warning (LDW) system that generates a vibration in the steering wheel when a road marking is crossed (e.g., Renault Clio 5 2019[4]). Therefore, if researchers or designers wish to use tactile feedback to inform drivers about the state of the LCA in a vehicle equipped with LDW, a different signal needs to be used, to avoid confusion between the two types of information. We tested a signal designed especially for LCA systems, using a metaphor of the system's action. When the LCA is activated, a slight jerk can be felt in the steering wheel as the system takes control of the steering column and the vehicle is centered in its lane. To induce an equivalent sensation for the drivers and appeal to their mental model, haptic feedback could take the form of a low-frequency vibration, or soft jerk. This jerk would be composed of two slight shakes that would remind them of the automated system taking control of the steering column. To avoid frightening the drivers, the jerk would have to occur without impacting the car's direction. As tactile feedback is usually brief, this jerk could be used to indicate transitions between LCA modes. Regarding the use of haptic feedback to indicate the LCA's current state, kinesthetic feedback seems more appropriate, given that this information can be presented continuously. As proposed by Katzourakis et al. (2014), a torque applied in the steering wheel can induce collaboration between the driver and the automated system. If the torque is sufficiently strong, resulting in a stiff steering wheel when the LCA is activated, it can inform the driver about the current state of the automated system. Finally, an important factor to consider when manipulating information in the steering wheel is the angle of the steering wheel. A jerk is presumably easier to perceive when the car is being driven along a straight road (i.e., when the driver is not exerting any force on the steering wheel) than when it is going round a bend (i.e., when the driver is exerting a force on the steering wheel). By contrast, a stiff steering wheel can only be felt when the driver exerts a force on it, which is more often the case in a bend than on a straight road.

Mainly used for LDW until now, haptic feedback has never been used in Level-2 vehicles for mode awareness purposes, that is, to indicate the current state of the LCA and mode transitions. Our general hypothesis was that mode awareness regarding the LCA system can be positively impacted by haptic feedback. A previous in-house study revealed that tactile feedback is easily perceived by drivers, but the settings for kinesthetic feedback have to be fine-tuned for it to be perceived. The present study had a twofold aim. In Experiment 1, we verified that kinesthetic feedback (transmitted through differences in stiffness in the steering wheel) is perceived by drivers. We expected to observe efficient differentiation between levels of kinesthetic feedback in a same/different task. In Experiment 2, we checked that tactile feedback (translated by a jerk

[4] See https://fr.e-guide.renault.com/fra/Clio-5/AIDE-AU-MAINTIEN-DE-VOIE

in the steering wheel), in combination with kinesthetic feedback, impacts mode awareness. We then examined the extent to which this awareness is moderated by the nature of the road (i.e., bends or straight). First, we expected the jerk and the stiff steering wheel to increase mode awareness. Second, we expected this effect to be modulated by the nature of the road, with the jerk being more efficient on a straight road than in a bend, and the stiffness being more efficient in a bend than on a straight road.

2. Experiment 1: Discrimination of Kinesthetic Feedback

This experiment was designed to assess whether kinesthetic feedback (i.e., differences in stiffness in the steering wheel) could be perceived by drivers in a driving simulator. To this end, we administered a same/different task. This task consists in presenting two stimuli one after the another and asking participants if they are identical or different (Serna et al., 2013).

2.1. Method

2.1.1. Participants

The sample was composed of 34 volunteers (10 women) aged 21–62 years (M = 45.68, SD = 10.95). They all held a driving license. They were recruited among employees at Renault Group's Technocentre site in France. They were not paid, and they all signed an informed consent form. Fourteen participants did not have a Level-2 vehicle, nine had driven a Level-2 vehicle before, one participant currently owned a Level-2 vehicle, and 11 had an LDW system or an LCA.

2.1.2. Driving Simulator and Scenario

Participants drove in a high-fidelity simulator, composed of the structure of a car and three TV screens occupying 145° of the driver's vision (see Figure 23). The simulator was fixed and used SCANeR software version 1.9 (AV Simulation, Boulogne, France, 2020) to simulate the driving environment. The simulated vehicle had an automatic gearbox. Participants could accelerate, brake, and turn the steering wheel. The instrument's cluster displayed the vehicle speed, but not the state of the automated systems (longitudinal and lateral control). Participants could only rely on the stiffness of the steering wheel to know when these systems were activated. The steering wheel was controlled by a SensoDrive electric motor system (SENSODRIVE, n.d.), which allowed haptic feedback to be provided by applying a torque and a vibration in the steering wheel.

Figure 23
Simulator used in Experiments 1 and 2.

The stiffness of the steering wheel varied according to the degree to which the automated lateral control centered the car in the lane. Three coefficients were specified in the simulator software: P, I, and D. Depending on the settings of these coefficients, the torque that the driver needed to exert to influence the direction of the car varied (see Table 24

Table 24 for detailed settings). P was the force error, and corresponded to the spring stiffness of the steering wheel. It was expressed in N m/°. I was the force of the integrated error and was expressed in N m/°/s. *D* was the force on the derivative of the error, and corresponded to the damping. It was also expressed in N m/°/s. Three levels of steering-wheel stiffness were used in the same/different task: no stiffness, moderately stiff, and very stiff.

Table 24

P, I, and D coefficient settings depending on steering wheel stiffness condition

Steering-wheel stiffness condition	P coefficient (N m/°)	I coefficient (N m/°/s)	D coefficient (N m/°/s)
No stiffness	0	0	0
Moderately stiff	0.05	0.0375	0.015
Very stiff	0.3	0.0375	0.015

2.1.3. Task

In the same/different task, participants to indicate whether two stiffnesses in the steering wheel, presented sequentially, were identical or different. Participants drove on a 2x2 divided highway with no other vehicles present. They had to activate the Level-2 systems and drive straight. Participants were informed that a level of stiffness had been introduced in the steering wheel. They were instructed to turn the steering wheel gently while staying in their lane to feel the stiffness. Shortly afterwards, participants were again informed that a level of stiffness had been introduced in the steering wheel. They were again instructed to turn the steering wheel to feel the stiffness, before answering the following question by the experimenter: “Are the stiffnesses different or identical?” The experimenter noted the answer, and the operation was repeated for each possible combination of stiffnesses.

2.1.4. Procedure

First, the participants were greeted and given a consent form. They were told that the goal of the experiment was to drive in a simulator with Level-2 driver-assistance systems and test a haptic interface in the steering wheel. The functioning of the Level-2 automated system was explained. Participants then seated themselves in the driving simulator. They drove on a straight road to familiarize themselves with the simulator, and were told that they could activate the Level-2 automated systems (lateral and longitudinal control). To do so, they had to inform the experimenter, who then activated the relevant system via a computer. When participants felt at ease with the simulator, we launched the task. Pairs of steering-wheel stiffnesses were presented sequentially. The order of the pairs was counterbalanced and randomly assigned to each participant. Each pair of stiffnesses was experienced once by each participant. The task lasted around 10 minutes. Once the task was completed, participants moved on to Experiment 2.

2.1.5. Experimental Design

The independent variable was stiffnesses pair. There were nine possible combinations of the three steering-wheel stiffnesses, resulting in nine experimental conditions: three pairs of *same* stiffnesses, and six pairs of *different* stiffnesses.

2.1.6. Measures and Statistical Analysis

Responses to the same/different task were orally pronounced by participants and recorded by the experimenter. Correct answers were scored 1, and incorrect ones 0. The means of correct response was calculated for each of the nine pairs. Participants could respond either *same* or *different*. Given that there was one correct response out of two possible answers, if the participants correctly perceived the differences or similarities within the each pair of stiffnesses, their correct response rate would be above 0.5. The statistical analysis consisted of a one-sample t test, comparing the mean correct response rate with a null hypothesis set at 0.5. The collected data did not respect normality of residuals for all the combinations of stimuli ($p < .001$). We therefore applied the Wilcoxon sum of rank test for each pair.

2.2. Results

For all six pairs of different steering-wheel stiffnesses (e.g., very vs. moderately stiff), participants' answers were above chance level (see Table 25). However, when the stimuli were identical (e.g., moderate vs. moderate), participants failed to identify that fact.

Table 25

Mean (Standard Deviation) and median (IQR) scores, and results of wilcoxon rank test for each steering-wheel stiffness pair.

Stiffness pairs	Mean (*SD*)	Median (IQR)	W	*p*
Very stiff – Very stiff	0.49 (0.51)	0.00 (1.00)	342	0.569
Very stiff - No stiffness	0.98 (0.16)	1.00 (0.00)	684	< .001
Very stiff - Moderately stiff	0.90 (0.32)	1.00 (0.00)	627	< .001
No stiffness – Very stiff	0.92 (0.28)	1.00 (0.00)	646	< .001
No stiffness - No stiffness	0.78 (0.42)	1.00 (0.00)	551	< .001
No stiffness - Moderately stiff	0.95 (0.30)	1.00 (0.00)	665	< .001
Moderately stiff – Very stiff	0.65 (0.48)	1.00 (1.00)	456	0.036
Moderately stiff - No stiffness	0.92 (0.28)	1.00 (0.00)	646	< .001
Moderately stiff - Moderately stiff	0.51 (0.51)	1.00 (1.00)	361	0.438

Note. H_a $\mu < 0.5$.

2.3. Discussion

The same/different task assessed whether participants were able to perceive and differentiate between the levels of stiffness in the steering wheel. Results revealed that participants were indeed able to distinguish between different stiffnesses, in line with the literature (Katzourakis et al., 2014), indicating that drivers are able to perceive sudden changes in LCA state in the form of kinesthetic feedback. This kinesthetic feedback was therefore used in Experiment 2. As the absence of stiffness in the steering wheel corresponded to manual driving, we only used the moderately stiff and very stiff levels in Experiment 2.

3. Experiment 2: Evaluation of Kinesthetic and Tactile Feedback

This experiment assessed the impact on mode awareness of a combination of tactile feedback (jerk in steering wheel) and kinesthetic feedback (stiffnesses tested in Experiment 1). It also aimed to evaluate the effect of the haptic feedback depending on the nature of the road. Participants drove on a single-lane road or divided highway in a driving simulator with Level-2 automated systems. They encountered LCA suspensions and had to detect them as quickly (reaction time) and accurately (SDT indices) as possible. We expected the combination of a jerk and a very stiff steering wheel to increase mode awareness (compared with no jerk and a moderately stiff steering wheel). We also expected this effect to be moderated by the nature of the road: the jerk would be more efficient on a straight road than in a bend, and the stiffness would be more efficient in a bend than on a straight road.

3.1. Method

3.1.1. Participants

The participants were the same as in Experiment 1. The volunteers were randomly assigned to one of the two steering-wheel jerk conditions (presence vs. absence).

3.1.2. Driving Simulator and Scenario

The simulator used in this experiment was the same as in Experiment 1. Based on the conclusions of Experiment 1, we used the *moderately stiff* and *very stiff* steering-wheel settings in this task (see Table 26 for a description of the degrees of stiffness). Moreover, a low-frequency vibration (jerk in the steering wheel) was used to indicate when the automated systems were suspended. This jerk was meant to be perceived as if the vehicle had driven over rails, in order to avoid confusion with other vibratory signals (e.g., high-frequency LDW vibrations). Different vibratory signals were used to signal the activation and suspension of the automated systems. A soft jerk indicated the activation, and a longer more important jerk indicated the suspension.

In this experiment, participants completed four driving scenarios: two on a 2x2 divided highway, and two on a single-lane road. The road was empty in all four scenarios. At the beginning of each scenario, Level-2 automation was activated, controlling both lateral and longitudinal movements of the car. Drivers kept their hands on the steering wheel and their attention on the road. Speed was regulated at 130 km/hr for the divided highway scenarios, and at 90 km/hr for the single-lane road scenarios. The scenarios were divided into four sections, each one corresponding to an event specific to the scenario. In the divided highway scenarios, the event corresponded to varying degrees of road marking erasure. In the single-lane road scenarios, the event corresponded to varying bend angles. For all scenarios, one of the four events resulted in the suspension of the LCA and the takeover of lateral control by the driver (see Table 3). Automated lateral control resumed shortly after the event, and automated longitudinal control remained activated throughout the scenario. It took around 5 minutes to complete each scenario. One event was encountered every 90 seconds, and lasted 10-20 s.

Table 26

Description of scenarios in detection task, depending on type of event and need for takeover.

Event	Divided highway A	Divided highway B	Single-lane road A	Single-lane road B
1	Road markings slightly erased	Road markings slightly erased	Right bend with slight angle	Right bend with slight angle
2	Road markings very erased: suspension of LCA	Road markings completely erased: suspension of LCA	Left bend with moderate angle	Left bend with slight angle
3	Road markings moderately erased	Road markings slightly erased	Right bend with moderate angle	Right bend with moderate angle
4	Road markings slightly erased	Road markings moderately erased	Right bend with sharp angle: suspension of LCA	Left bend with sharp angle: suspension of LCA

3.1.3. Task

The goal of the detection task was to detect when the LCA suspended, and indicate this by pressing a button located below the right lever of the steering wheel. Participants drove in two different types of scenarios: a divided highway and a single-lane road. Each scenario was divided into four sections. For example, in the divided highway scenarios, each section corresponded to a different degree of lane marking erasure. In one of the four sections, a Level-2 automated system was suspended, giving lateral control of the car back to the driver. Longitudinal automated control of the vehicle remained activated throughout. When the drivers felt that they had taken over the lateral control, they had to press a button located on the steering wheel as quickly as possible. In the single-lane road scenario, the drivers had to negotiate four bends of different angles. In the one with the sharpest curve, the Level-2 automated lateral control was suspended. Participants then had to press the button on the steering wheel as quickly as possible.

3.1.4. Procedure

The detection task directly followed the same/different task. It was composed of four driving scenarios, each lasting for 5 minutes, for a total of approximately 20 minutes. The order of the scenarios was counterbalanced across conditions and randomly assigned to participants. Once the task was completed, the participants were interviewed, rated items that will not be discussed in this paper, and completed a sociodemographic questionnaire. They were then thanked for their participation.

3.1.5. Experimental Design

A 2 (between-participants) × 2 (within-participants) × 2 (within-participants) experimental mixed design. The between-participants factor was the presence of a jerk in the steering wheel when the LCA suspended. It had two modalities (jerk present vs. absent). The first within-participants factor was the stiffness of the steering wheel when the LCA was activated. It had two modalities (moderately vs. very stiff steering wheel). The second within-participants factor was the type of scenario. It had two modalities (single-lane road vs. divided highway).

3.1.6. Measures and Analysis

To study the effects of the haptic feedback while driving with Level-2 automated systems, binary data were collected by the simulation software. These data represented the state of the automated systems (binary) and participants' reaction times (timestamped button presses). Two measures were extracted from these data: signal detection indices and reaction times.

Signal Detection Index Calculation

Signal detection indices were calculated based on the responses to the detection task. Steering wheel behaviour was the only information given by the interface to inform on the state of the LCA. It was possible for participants to mistakenly press the button, thinking that the automation had suspended when it was not the case. It was also possible for them not to detect the takeover, and not to press the button when they needed to. According to SDT, there are four different categories of action: hit, miss, correct rejection, and false alarm (Stanislaw & Todorov, 1999). The classification of each action depends on the response in the presence or absence of the stimulus. In this study, the presence and absence of the stimulus were represented as the activation and suspension of Level-2 automated lateral control. The action was represented by participants pressing the response button (see Table 27).

Table 27

Correspondence between driver's button presses and current state, according to sdt action categories.

	Button pressed	Button not pressed
LCA suspended	Hit	Miss
LCA activated	False alarm	Correct rejection

Based on the binary data and the timestamps collected with SCANeR software, we were able to see whether participants had pressed the response button or not, depending on the state of the vehicle's automated lateral control. The event on the road (e.g., bend in the single-lane road scenario) occurred approximately at the same point for every participant, plus or minus 10 seconds. We were therefore able to establish time windows for each of the events in the scenario during which pressing the response button was deemed to be a response. No responses outside these time windows were considered. For the single-lane road scenario, the time window was set at 40 seconds: 10 seconds before the bend, 20 seconds during the bend, and 10 seconds after the bend. For the divided highway scenario, the time windows were set at 37 seconds: 10 seconds before the area where the road markings were erased, 17 seconds during which road markings were erased, and 10 seconds after the road markings reappeared. For both scenarios, the first press on the response button during the time window was deemed to be a response. During the time windows concomitant with the LCA suspension, a response was considered to be a *hit*, and no response was considered to be a *miss*. During time windows concomitant with no LCA suspension, a response was classified as a *false alarm* (FA), and no response as a *correct rejection* (CR). In line with SDT, we calculated *d'* and β indices (see Stanislaw & Todorov, 1999, for descriptions of index calculation). The *d'* value represents the ability of participants to discriminate between a stimulus (here, suspension of automation) and noise (here, automation still activated). *d'* usually varies between 0, representing random responses, and 4.65, representing near-perfect discrimination between signal and noise. β indicates the response bias (here, tendency of participants to respond more often that automated systems were either active or inactive). A positive β indicates that participants tend to be *conservative* (here, estimation that LCA was more often active). A negative β indicates that participants tend to be *liberal* (here, estimation that LCA was more often inactive) (McNicol, 1972).

Reaction Times

The data (i.e., timestamp, LCA state, and status of button that drivers had to press) were collected from the SCANeR software at a sampling rate of 60 Hz. The state of the system and the status of the button were both binary: 0 meant that the system was suspended and the button was not pressed; and 1 meant that the system was activated and the button was pressed. To calculate reaction times, two timestamps were identified: one when the system suspended, and the other when the participant first pressed the button. The period between these two timestamps was deemed to be the time taken by participants to detect the suspension of the automated system. If participants did not press the button or pressed the button more than 20 seconds after the suspension, the reaction time was not taken into account. Reaction time measurements were transformed using a logarithmic function to correct the positive skewedness that is usually observed with reaction times (Howell, 2012). When we tested assumptions of homogeneity of variance and normality of residuals (i.e., Shapiro-Wilks), results were nonsignificant (all *p*s > .05). We therefore ran a mixed analysis of variance (ANOVA) on three factors: presence of jerk, steering-wheel stiffness, and scenario. Tukey's post hoc tests were performed on those interactions that were statistically significant.

3.2. Results

3.2.1. Signal Detection

The presence of the jerk seemed to have an impact on the detection of mode transition, which was better when the jerk was present (*d'* = 2.85) than when it was absent (*d'* = 1.90; see Table 28 for a summary). Participants in the absent jerk condition tended to be more conservative (β = 1.57) and to press the button less often than those with the jerk (β = 0.67). Regarding the stiffness of the steering wheel, the moderately stiff steering wheel induced slightly better detection of mode transition (*d'* = 2.41) than the very stiff steering wheel did (*d'* = 2.23). Participants tended to be conservative in both conditions, but participants were less inclined to press the button in the moderate stiffness one (β = 1.81 vs. β = 1.40). Finally, the detection of LCA suspension was more accurate (*d'* = 2.48) , in the divided highway scenario than in the single-lane road scenario (*d'* = 2.11). Participants tended to be less conservative in the former (β = 1.32) than in the latter (β = 1.72).

Table 28

Hit Rate, FA Rate, d' and β depending on each manipulated factors.

Factors	Hit rate (SD)	FA rate (SD)	d'	*β*
Jerk				
Present	0.94 (0.24)	0.10 (0.22)	2.85	0.67
Absent	0.88 (0.32)	0.24 (0.36)	1.90	1.57
Stiffness				
Moderate	0.93 (0.26)	0.17 (0.32)	2.41	1.81
Very	0.90 (0.31)	0.17 (0.30)	2.23	1.40
Scenario				
Single-lane road	0.91 (0.29)	0.21 (0.33)	2.11	1.72
Divided highway	0.91 (0.29)	0.13 (0.27)	2.48	1.32

3.2.2. Reaction Times

Reaction time was measured for participants who correctly detected LCA suspensions. Thirteen participants did not detect mode transitions in any of the conditions, and were excluded from the analysis. Analysis of the time between LCA suspension and the pressing of the button by participants revealed differences in reaction times according to the factors. The first ANOVA revealed a significant effect of scenario on reaction time, $F(1, 21) = 7.67$, $p = .012$, $\eta^2p = .27$ (see Table 29). Participants reacted faster in the divided highway scenario ($M = 2.97$; $SD = 1.68$) than in the single-lane road scenario ($M = 4.12$, $SD = 2.7$). Analysis also revealed a significant effect of stiffness, $F(1, 21) = 10.27$, $p = .004$, $\eta^2p = .33$. When the steering wheel was very stiff ($M = 3.20$, $SD = 1.75$), participants responded faster than when it was only moderately stiff ($M = 3.87$, $SD = 2.52$). No significant main effect of the presence or absence of the jerk was found ($p > .05$).

Table 29

Results of ANOVA according to the experimental conditions.

	F	p value	η^2_p
Scenario	7.663	0.012	0.267
Stiffness	10.272	0.004	0.328
Jerk	0.183	0.673	0.009
Scenario*Stiffness	0.614	0.442	0.028
Scenario*Jerk	6.955	0.015	0.249
Stiffness*Jerk	8.941	0.007	0.299
Scenario*Stiffness*Jerk	0.897	0.354	0.041

There was an interaction effect between jerk and scenario, $F(1, 21) = 6.96$, $p = .015$, $\eta^2p = .25$. The presence of the jerk induced a faster reaction time in the divided highway scenario ($M = 2.53$, $SD = 1.05$) than in the single-lane road scenario ($M = 3.39$, $SD = 2.06$; see). This difference was confirmed by Tukey's post hoc test, $t(21) = 3.37$, $p = .049$). None of the other post hoc tests for this interaction were significant (all ps > .05). The presence of the jerk also interacted with the stiffness in the steering wheel, $F(1, 21) = 8.94$, $p = .007$, $\eta^2p = .30$. When the jerk was absent, a very stiff steering wheel induced a significantly faster reaction time ($M = 2.86$, $SD = 1.61$) than a moderately stiff one did ($M = 4.46$, $SD = 3.21$), $t(21) = 4.21$, $p = .003$

(see Figure 25). None of the other post hoc tests were significant for this interaction (all *p*s > .05).

Figure 24

Detection task reaction time according to presence/absence of jerk and scenario.

** p < .05.*

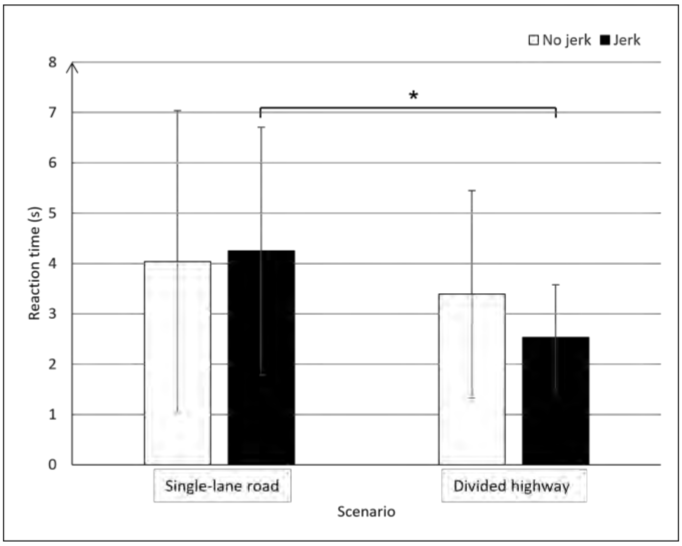
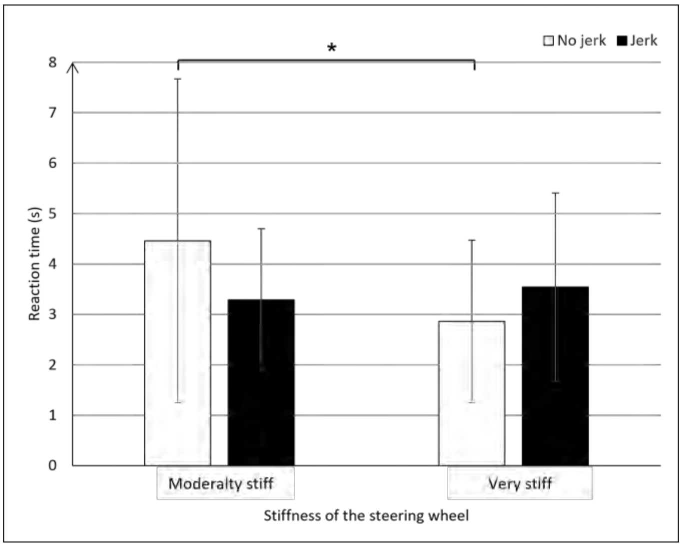


Figure 25

Detection task reaction time according to presence/absence of jerk and steering-wheel stiffness.

** p < .05.*



4. General discussion

The present study assessed whether kinesthetic feedback (i.e., steering-wheel stiffness) and tactile feedback (i.e., jerk in steering wheel) enhanced mode awareness in Level-2 automated vehicles in road bends or on straight roads. Mode transitions were detected faster with kinesthetic feedback. The presence of tactile feedback in the steering wheel induced more accurate detection of mode transition. When detection was correct, no effect of tactile feedback, either negative or positive, was observed on reaction time. However, the effect of the tactile feedback was modulated by type of road and kinesthetic feedback.

The feedback seemed to have a beneficial impact on mode awareness in situations where the LCA system was suspended. Regarding tactile feedback, in the form of a very low-frequency vibration, it appeared to induce better detection of mode transition. This result is in line with the findings of Suzuki and Jansson (2003) for LDW systems. These authors found that a vibration in the steering wheel enabled participants to efficiently correct the trajectory of the vehicle after a deviation. One main advantage that these authors shed light on was the ability of drivers to efficiently understand the haptic feedback without training. The participants' mental models facilitated the understanding of the message, as information transmitted by the steering wheel most likely meant that an event was taking place that was related to the trajectory. A similar phenomenon must have occurred in Experiment 2, as participants understood that the tactile feedback in the steering wheel was related to an event concerning the LCA's state, allowing them to efficiently detect the transition with minimal training. This kind of feedback appears to be appropriate for informing drivers about transitions related to LCA. However, we failed to observe any positive impact of tactile feedback on reaction times following LCA suspension. Reaction times were no longer with versus without the tactile feedback. It is worth noting that reaction times with the tactile feedback were shorter on the divided highway than on the single-lane road. There are two possible explanations for this result. In the single-lane road scenario, the LCA suspension took place in a bend. Participants may therefore have had to correct the trajectory of the vehicle to avoid a hazardous situation before pressing the button on the steering wheel. In the divided highway scenario, the LCA suspension took place on a very slight bend. Drivers were not faced with any immediate danger, and no urgent maneuver was necessary. They were therefore free to press the button and were able to do so more quickly. An alternative way of measuring reaction time in the bend would have been to calculate the interval between the LCA suspension and the return of the vehicle to its lane after controlling the direction. Another possible explanation is that when a car is going round a bend, the steering wheel is turned and force applied to it by the driver, meaning that any tactile feedback provided at that point is less perceptible than it would be if the car were travelling along a straight road, with minimal force exerted by the driver.

The kinesthetic feedback also significantly impacted mode awareness. The detection of mode transitions was slightly impacted by the kinesthetic feedback, as drivers identified LCA suspensions more efficiently with the moderate kinesthetic feedback. However, when the detection was correct, the very strong kinesthetic feedback induced a faster reaction time than the moderate kinesthetic feedback did. These results confirm the suggestions of Katzourakis et al. (2014) about the effect of torque. These authors concluded that haptic feedback in the form of torque resulting in a stiff steering wheel only prevents lane deviation if the steering is highly automated. Our results revealed that kinesthetic feedback in the form of a stiff steering wheel reduced the amount of time needed to detect a takeover when an automated steering system was suspended. Interestingly, the effect of stiffness was modulated by the presence of tactile feedback. When the latter was absent, reaction times were faster with the very strong versus moderate kinesthetic feedback. The redundancy of haptic messages resulting from the combination of tactile and kinesthetic modalities did not improve reaction times, but did not decrease them either. This may indicate that the demands on the tactile sensory channel nearly overwhelmed participants' processing capacity. According to Wickens' multiple resources model (2008), if one sensory channel is burdened by too many demands, it results in mental overload and poorer performance. Our results suggest that if any more information had been conveyed through the haptic modality, performance would have declined and reaction time increased. More studies are therefore needed to quantify the amount of information that can be transmitted through the haptic channel while driving.

Before testing the kinesthetic feedback steering wheel in a driving task, we ensured that the stiffness settings were sufficiently different to be perceived by the drivers. The result of Experiment 1 showed that participants could efficiently distinguish between the levels of kinesthetic feedback we provided. We were therefore able to use the kinesthetic feedback in the steering wheel in Experiment 2. This two-step method followed of situational awareness model of Endsley (1995). According to this model, it takes three steps of information processing to achieve optimum situational awareness and, by extension, mode awareness: perception, comprehension, and projection into the future. In Experiment 1, we assessed whether participants correctly perceived the kinesthetic feedback. In Experiment 2, we assessed whether the information was comprehended and allowed participants to project the future states of the automated system. This method could be used in future research on haptic feedback and its impact on mode awareness. It could also be used to explore feedback in different sensory modalities, such as auditory feedback (Monsaingeon et al., 2021).

4.1. Limitations

The two experiments had several limitations. First, some data losses were caused by the fact that each experimental condition was encountered only once in Experiment 2. If participants did not press the button when required, no reaction time was recorded. This meant that we could not include some participants in the analysis of reaction times. With more repetitions of the experimental conditions, the number of participants included in the analysis would have been greater. Second, the simulator was static, and only the haptic interface allowed participants to ascertain the state of the LCA. In a real vehicle, visual interfaces are present, but most importantly, proprioceptive sensations of the car's movement o can indicate when the LCA has stopped controlling the vehicle. A high-fidelity simulator with a moving cabin would allow the effect of haptic feedback to be assessed in a more realistic way.

4.2. Conclusions and Recommendations

To conclude, this study revealed a positive effect of haptic feedback on mode awareness. Both tactile and kinesthetic types of haptic feedback can be used as signals to inform drivers of the state of Level-2 automated systems. Tactile feedback, in the form of a jerk in the steering wheel, can be used as a brief and discrete message about an LCA mode transition. Kinesthetic feedback, in the form of a stiffer steering wheel when Level 2 is activated, can be used as a continuous message that LCA is currently operating. These messages can be integrated into vehicles that already possess haptic feedback, such as vibrations for LDW. However, further studies manipulating the presence of all these haptic messages at the same time are required. If all messages are transmitted at the same time through the haptic channel, this may trigger cognitive overload. Multimodal interfaces using other modalities (e.g., visual or auditory) could avoid cognitive overload by allowing drivers to find the information they need in the different sensory channels available to them. Auditory or visual information could be tested with regard to its ability to efficiently induce mode awareness, as depicted in the model of

Endsley (1995). It would therefore be relevant for future studies to investigate the effect of the tested haptic feedback in combination with interfaces using other sensory modalities.

Points clés

- Le volant peut fournir deux types de signaux haptiques : tactile (perception de la peau) et kinesthésique (effort musculaire).

- Différentes intensités de signaux kinesthésiques dans le volant peuvent être perçues et différenciées.

- La présence de signaux tactiles induit une discrimination efficace des suspensions des systèmes automatisés.

- Les retours tactiles et kinesthésiques interagissent avec le temps de réaction sur la détection d'une suspension de systèmes automatisés.

Key points

- Two types of haptic feedback can be provided in the steering wheel: tactile (perception from the skin) and kinesthetic (muscular effort)

- Different degrees of kinesthetic feedback in the steering wheel can be perceived and differentiated

- The presence of a tactile feedback induces efficient discrimination of suspensions of automated driving systems

- Tactile and kinesthetic feedback interact on reaction time to detect a suspension of automated driving systems

Chapter 9 – Multimodal interface and reliability displays: effect on attention, mode awareness and trust in partially automated vehicles

This chapter aims to combine multimodal interface with a reliability display. All the unimodal interfaces tested in the Experimental section were improved based on the results of the studies and were gathered in a single multimodal interface. A longitudinal study evaluating the capacity of this interface to stimulate appropriate level of attention, mode awareness, and trust in automation is described in this Chapter. This study allowed to highlight the importance the effect of symbols introduced in the IPLA on attention allocation. It allowed to evaluate its effect on mental models and driving behaviours, and the arising trust in automation. The experimental study presented in this chapter is the subject of a research article in preparation. The introductory elements of this article, already mentioned in the previous chapters, have been reduced to avoid redundancies with previous chapters.

Monsaingeon, N., Caroux, L., Langlois, S., & Lemercier, C. (In preparation). Multimodal interface and reliability displays: effect on attention, mode awareness and trust in partially automated vehicles

Résumé

L'objectif de cette étude est d'évaluer l'effet d'une interface multimodale indiquant les limites des automatisations sur l’allocation de l’attention, la conscience des modes, et la confiance dans l'automatisation. Les participants ont conduit dans un simulateur de conduite avec des systèmes d’automatisation partielle de la conduite et ont été confrontés à des situations suspension de ces systèmes dans différents contextes. Ils ont conduit pendant trois sessions de conduite, avec soit une interface multimodale indiquant les limites de l'automatisation, soit une interface visuelle classique. Leurs performances de conduite lors des suspensions, leurs comportements oculaires, leur confiance dans l’automation, et leur charge mentale ont été évalués. Les résultats ont révélé que l'interface multimodale stimulait l'attention à manière appropriée, augmentait la conscience du mode et la confiance dans l'automatisation, mais que ces effets dépendaient des situations de conduite. Les indications sur les limites de l'automatisation ont amélioré la connaissance de l'automatisation, mais cette connaissance n'a pas nécessairement conduit à une amélioration des performances de conduite. Des solutions de conception sont discutées pour favoriser l'amélioration des performances de conduite.

Abstract

The goal of this study is to evaluate the effect of a multimodal interface indicating the limits of automation to stimulate appropriate level of attention, induce accurate mode awareness and trust in automation. Participants drove in a driving simulator with partially automated systems and were confronted with surprising situations of suspension of automated systems in different contexts. They drove the simulator during three driving sessions, with either a multimodal interface indicating limits of automation or a visual basic interface. Their driving performances, ocular behaviours, subjective evaluations of trust and workload were evaluated. The results revealed that the multimodal interface stimulated appropriated level attention, increased mode awareness and trust in automation, but that these effects were context dependent. The indications of the limits of automation improved the knowledge regarding automation but this knowledge did not necessarily lead to improved driving performances. Design solutions are discussed to support improvement of driving performances.

1. Introduction

Automated systems can supervise lateral and longitudinal controls of the vehicle thanks to a combined function approach (NHTSA, 2013). In situations such as sharp bends, automated longitudinal system can reach its limits, giving control of the direction back to the driver. When a limit of automated systems is reached, the driver needs to be ready to control the vehicle, be aware of the state of automation. Trust that drivers have in automated systems is correlated with attention allocation (de Winter et al., 2014). It was proposed by previous authors that improving knowledge about automation's limits lead to improving trust in automation (Seppelt & Lee, 2019). Interface of automated vehicles need to follow guidelines to address these challenges: stimulate appropriate level of attention and intervention, provide required understanding of automation's capabilities and states, minimize automation surprises, and provoke adequate calibration of trust (Carsten & Martens, 2019). Multimodal interfaces using auditory and haptic feedback offers the possibility to convey information without having the driver to gaze at the instrument's cluster, stimulating appropriate level of attention and informing on changes of modes of automated systems. Interfaces indicating limits of automated systems allow the driver to anticipate changes of modes of automated systems and place adequate trust in the system. This study proposed to investigate the effect of prolonged usage of a multimodal interface indicating the limits of automation on the stimulation of attention, mode awareness and trust in automated systems.

1.1. Challenges of Automated Driving

Goals and guidelines for interfaces of automated vehicles were established by Carsten and Martens (2019) to respond to the main challenges raised by automated driving. Those goals are based on Rasmussen's (1983) SRK model of performances of skilled human operators. Six goals have been identified:

« (1) Provide required understanding of the automated vehicles capabilities and status (minimise mode errors); (2) Engender correct calibration of trust; (3) Stimulate appropriate level of attention and intervention; (4) Minimise automation surprises; (5) Provide comfort to the human user, i.e. reduce uncertainty and stress; (6) Be usable. » (Carsten et Martens, 2019, p. 5).

Goals (1) provide required understanding of automated vehicles capabilities and status, and (4) minimize automation surprises, can be gathered under a common global goal of increasing mode awareness, as they both refer to the correct perception, comprehension, and projection of automation's mode (Endsley, 1995). In this study, we evaluated the effect of interfaces on the following challenges: Stimulate appropriate level of attention and intervention, increase mode awareness, and engender correct calibration of trust.

1.1.1. Stimulate Appropriate Level of Attention and Intervention

A first challenge that comes across drivers when using partially automated vehicles is to correctly distribute their attentional resources. They are usually allocated to two activities, depending on the state of automation: monitoring the actions of automation, and controlling the vehicle (Carsten & Martens, 2019). Situations in which automation suspends suddenly, for example when passing a bend road with too elevated speed, require the attentional resources to be allocated to controlling the vehicle. The drivers need to rapidly comprehend the new mode of automation and that a takeover of the direction of the car is required.

1.1.2. Induce Accurate Mode Awareness

Experience with automated system plays a major role in correct understanding of the role one has to play in the interaction (Solís-Marcos et al., 2018). Interactions with an automated driving system allow to forge a representation of its purpose, form, functioning, state and structure, which can be merged into the term *mental model* (Seppelt & Victor, 2020). The more the users interact with an automated driving system, the more precise their mental model will be (Beggiato et al., 2015; Forster et al., 2019). Therefore, longitudinal studies are necessary to capture the evolution of mental models. As proposed by Kurpiers et al. (2020), the assessment of mode awareness can be performed by measuring three dimensions: the driving behaviour of the drivers, their ocular behaviour, and their mental models. The behaviour of the driver should be adapted to the mode of automation. When switching to manual driving, Deviation from Central Lane (DCL) or Time Headway (TH) should reveal that the drivers are in control of the vehicle. The ocular behaviour of the drivers should reveal that their gaze is fixed on the exterior of the environment when a takeover occurs. The mental models, evaluated through questionnaires regarding the functioning of automation depending on the situation, should be accurate.

1.1.3. Induce Adequate Trust Calibration

Trust in automation is impacted by the way the drivers perceive and understand its functioning. If automation does not accomplish the goals that it is meant to achieve, breakdown of trust can be observed (Parasuraman & Riley, 1997). With highly automated vehicles, drivers over trusting the automated system gazed less at the road (de Winter et al., 2014) or failed to take over correctly when needed (F. O. Flemisch et al., 2014). Therefore, trust in automation should be appropriately calibrated to the automated system's capacity and limits. Informing the drivers of the automated systems' capacity and reliability allow them to place adequate trust in it (Helldin et al., 2013).

1.2. Multimodal Interfaces and Reliability Information

To address the challenges mentioned above, two types interface were investigated in this study: multimodal interface and interface of reliability of automation.

1.2.1. Multimodality

Multimodal interfaces use multiple sensory modalities to convey information. It allows to distribute the attentional demands of the interface on the multiple sensory channels, reducing the cognitive load compared to if all demands were directed to only one sensory channel (Wickens, 2008). Earcons indicating transitions of control from the system to the driver allow them to re-engage in the driving task (Petermeijer et al., 2017). Haptic feedback, in the form of kinesthetic and tactile feedback in the steering wheel, induce quick response and are easily understood (Murata & Kuroda, 2015). Haptic feedback should reduce the risk of mode errors by informing the driver of the state of automation efficiently. Altogether, multimodal interfaces should induce more appropriate repartition of visual attention and make the identification of modes more accessible, allowing to respond to the first two challenges posed by automation: stimulate appropriate level of attention and induce accurate mode awareness. In addition to multimodal interfaces, information regarding the reliability of automation should allow the drivers to anticipate transitions of modes.

1.2.2. Reliability Displays

Interfaces indicating the limits of automation should allow the drivers to direct their attentional resources to the road when control is needed, addressing the first challenge of automation: stimulate appropriate attention. However, the addition of visual information can capture attention cause mental workload (Monsaingeon et al., 2019). It can be expected that with training to use the interface, mental workload would decrease (Christoffersen et al., 1996). By informing on the limits of automation, drivers should be able to anticipate transitions of modes and learn to identify situations in which these transitions can occur, addressing the second challenge of automation: induce accurate mode awareness. Finally, by informing on the situations that automation can or cannot deal with, drivers should calibrate their trust accordingly, addressing to the third challenge of automation.

1.3. Methodology Overview

The experimental method of this study aimed to assess the longitudinal effect of interface modalities on attention allocation, mode awareness and trust in automation while interacting with partially automated systems. Participants were recruited following specific criteria. They were assigned to either one of two interface conditions. Participants were prepared for the study with educational material. Then, in a driving simulator, participants drove with Level-2 automated systems in driving scenarios built for the purpose of this study. The scenarios depicted driving situations in which automation could suspend depending on environmental conditions. These use cases were selected with experts of the automotive industry because of their representativeness of current automated systems functioning. The use cases were bent roads, erased road markings, traffic jams and foggy areas. For three weeks, the participants performed 6 driving sessions, the first and last one being considered as pretests and post-tests, and the four in-between being considered as training. Along the study, participants' driving behaviours, visual fixations were measured, their mental models, trust in automation, and workload were rated. The analysis plan consisted in comparing the measures of participants between the pretests and post-tests, depending on their interface condition, and separately for each use case.

1.4. Research Objectives

This study aimed to evaluate to what extent the prolonged exposition to a Multimodal Interface with indicator of Limits of Automation (MILA) addresses the goals proposed by Carsten and Martens (2019). The tested goals were to stimulate appropriate level of attention and intervention, to induce accurate mode awareness, and to induce appropriate trust in the system. A driving simulator experiment took place. The goals of Carsten and Martens were turned to general hypothesis and divided into operational hypothesis: (1) a MILA interface stimulates a more appropriate level of attention than a Visual Basic Interface (VBI), which would translate in more gaze fixations on the instrument's cluster before the suspension of automated systems, more important mental workload for MILA than VBI on first usages but a decrease after multiple driving sessions; (2) MILA induces a more accurate mode awareness than VBI which would translate in more precise mental models in shorter periods of time for MILA than for VBI, better control of the vehicle when automated systems suspend will be better for MILA than for VBI, more important visual fixations on the road when automation suspends for MILA than for VBI; (3) MILA induces a more important trust in automation than VBI.

2. Method

2.1. Participants

The sample was composed of 40 volunteers (15 women) aged 39–65 years (M = 53.34, SD = 6.83). They were recruited via the panelist Eurosyn. It was required to be able to drive without glasses, to hold a valid driving license for at least 3 years, to drive several times a week, to have experienced cruise control at least once, and to have a positive attitude toward automation (evaluated on a Likert-style rating scale). If volunteers met these requirements, they were tested on their crystallized and fluid intelligence, and their visual acuity. The crystallized intelligence was assessed with the WAIS-IV's Vocabulary test. Volunteers had to define concepts and objects (e.g., a mirror) and were evaluated on the quality of their definition. Seven volunteers were excluded because they failed to define three consecutive concepts (M = 31.8; SD = 6.03). Fluid intelligence was assessed with the WAIS-IV's Cancellation test. Volunteers had to cross out targets among distractors in a limited time. The score considered the speed of execution, the number of correct and incorrect responses (M = 15.98; SD = 12.84). Then, participants took a visual acuity test in which they had to read a text with small letters at 60 cm of distance. Their a priori trust in automation was assessed with a Likert-style rating scale. Participants were randomly assigned to one of two interface conditions. The a priori trust in automation did not differ significantly between the two interface groups ($p < .05$). Participants signed an informed consent form and were paid 150 euros for their participation. The majority of the participants had a cruise control in their vehicle ($n = 27$), some had an ACC ($n = 12$), and a few did not have either a cruise control or an ACC, but already used it ($n = 2$). Sixteen participants reported using their cruise control as much as possible, thirteen use it when the situation seems appropriate, 4 reported using it sometime and 8 never use it. Prior to the experiment, participants were explained the principles of partially automated driving, functioning and situations to use it.

2.2. Material

2.2.1. Driving Simulator

A high-fidelity driving simulator was built for the purpose of CMI Project at IRT SystemX (Palaiseau, France) where this study took place (see Figure 26). It was composed of a full-car cab with seven visual channels, providing a high-fidelity graphic resolution and realistic driving environment. Three visual channels were located in front of the vehicle providing a 180° field of view. Three visual channels were display screens showing the view from rear-view and side mirrors. The remaining visual channel was a virtual instrument's cluster displaying the instrument cluster. SCANeR software version 1.9 (AV Simulation, 2020) was used to simulate the driving environment. The simulated vehicle had an automatic gearbox and two modes of automated driving could be activated. The steering wheel was controlled by a SensoDrive electric motor system (SENSODRIVE, n.d.), which allowed to produce haptic feedback by applying a torque and a vibration in the steering wheel. Auditory signals in the form of earcons were emitted from the driver’s headrest.

Figure 26

Driving simulator composed of a car with a 180° field of view.

2.2.2. Automated Systems

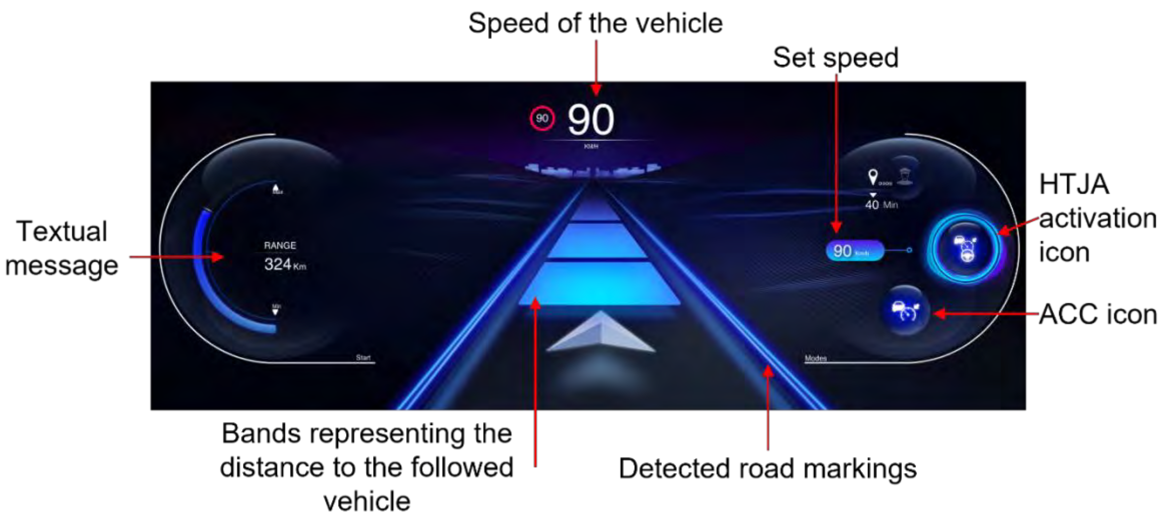
Automated driving systems were integrated in the simulator to simulate the automated systems available in today's Level-2 vehicles. Two modes of automated driving were simulated: the ACC and the Level-2 automated system called Highway and Traffic Jam Assist (HTJA) that was composed of a LCA and an ACC. They could be activated by pressing buttons of a tactile screen on the right part of the steering wheel's arm. Drivers could change mode as they wished, switch from manual to ACC and HTJA. Two limits of the ACC could be reached: maximum deceleration and non-detection of the followed vehicle. The maximum deceleration limit was reached when approaching a slow vehicle. The non-detection of the followed vehicle was reached when fog blocked the sensors. There were two limits of the HJTA: maximum lateral acceleration and non-detection of road markings. The limit of lateral acceleration could be reached when passing sharp bends and reaching an important lateral acceleration. The limit of detection of road markings was reached when the road markings were importantly or fully erased. When a limit of HTJA was reached, the LCA suspended while the ACC remained active. Once the correct condition for the LCA were present, it became active again on its own.

2.2.3. Interface Design

Two interfaces were compared during this study: a VBI and a MILA. These two interfaces shared similarities. They both presented the speed of the vehicle, the set speed of the ACC, the current state of the HTJA and of the ACC, the detected road markings, the set distance of the ACC and a textual message area (see Figure 27). When the distance with the lead vehicle was too close with a TTC under 4 seconds, an auditory and visual alert was emitted. The alert was played again if TTC was below 2s. The two interfaces differed in the information they transmitted regarding the state and functioning of automated systems. The VBI only displayed the states of automated systems on the instrument's cluster. It was also the case of the MILA, with the addition of an indicator of limits of automation, a haptic interface, and an auditory interface.

Figure 27

Visual representation of the instrument's cluster of the VBI, with information that was mutual to both interfaces.



Indicator of Proximity to the Limits of Automation

An IPLA was presented in this interface when HTJA was activated. The IPLA informed the drivers of a risk of a transition of state of the HTJA with the objective of allowing them to anticipate the transition and act appropriately. Its design was based on a prior study and has been improved to reduce the risk of inadequate behaviours (Monsaingeon et al., 2021). It was displayed on the instrument's cluster in order to be perceived in peripheral vision. The limits of two systems of the HTJA were displayed: limits of LCA and limits of ACC which also provoked a suspension of LCA. In both representations, a cloud was displayed with varying size depending on the proximity to the limits of automation (see Table 30). Two degrees of limits were indicated: a moderate size and yellow cloud indicated that limits are getting closer but should not be reached, a large red cloud indicated that limits will soon be reached. A pop-up screen appeared in the centre of the instrument's cluster when the yellow and red clouds were displayed. It represented a visual icon of the event that caused the approach of limits (e.g., representation of a bent road), a textual message of the cause of approach to the limits and the action to perform to act appropriately.

Auditory interface

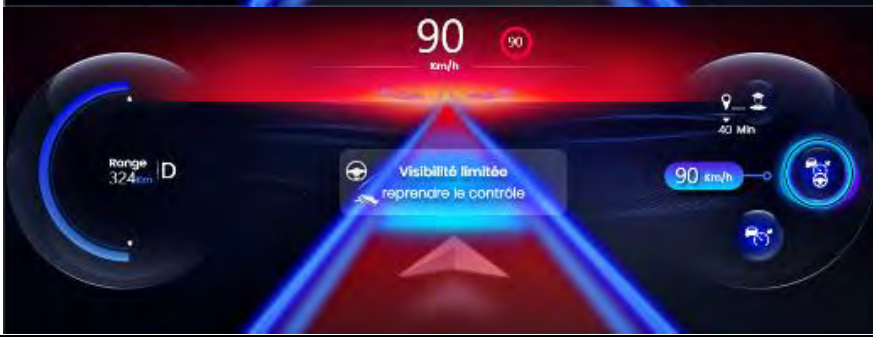
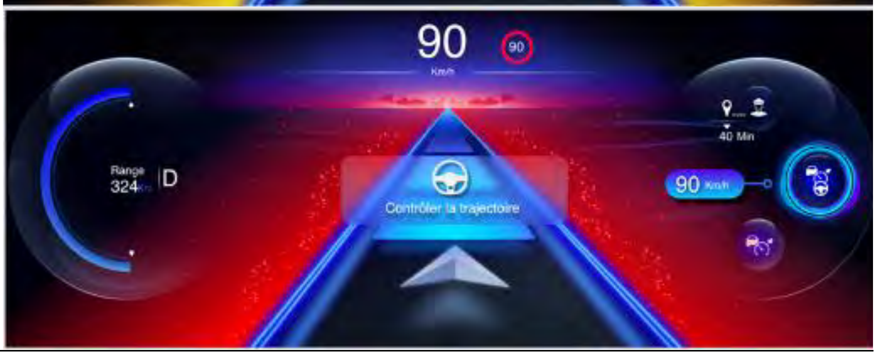
The auditory interface indicated when a transition of control from the system to the driver occurred. It was composed of two earcons. The efficiency of the earcons to be perceived and comprehended were evaluated in previous studies. One earcon was presented when HJTA transited to ACC only, meaning that it indicated a transition of control of lateral movements of the car from the system to the driver. It was composed of two descending notes. A second earcon was presented when the HTJA (ACC and LCA) suspended, indicating a transition of control of both lateral and longitudinal movements of the car from the system to the driver. It was composed of three descending notes. The earcons were validated through different experiments to ensure that they were perceived and comprehended (Monsaingeon et al., 2021).

Haptic Interface in the Steering Wheel

Two haptic signals were transmitted through the steering wheel: kinesthetic and tactile signals. The settings of the haptic interface were tuned after the results of inter studies. The kinesthetic signal consisted in increasing the stiffness of the steering wheel when HTJA was activated. The tactile signals consisted in indicating a transition of control of lateral movement of the car through a low frequency vibration in the steering wheel. Two soft jerks indicated the activation of HTJA, three moderate jerks indicated a suspension of HJTA. The correct perception and utility of the haptic interface validated in previous experiments of Chapter 8.

Table 30

Representations of the IPLA depending on the proximity to the limits and the automated system.

Type of limit and degree of proximity to the limits	Representation of proximity to the limits of automation
Limit of ACC	
Limits at moderate proximity	
Limits at close proximity	
Limit of LCA	
Limits at moderate proximity	
Limits at close proximity	

2.2.4. Eye-tracking Glasses

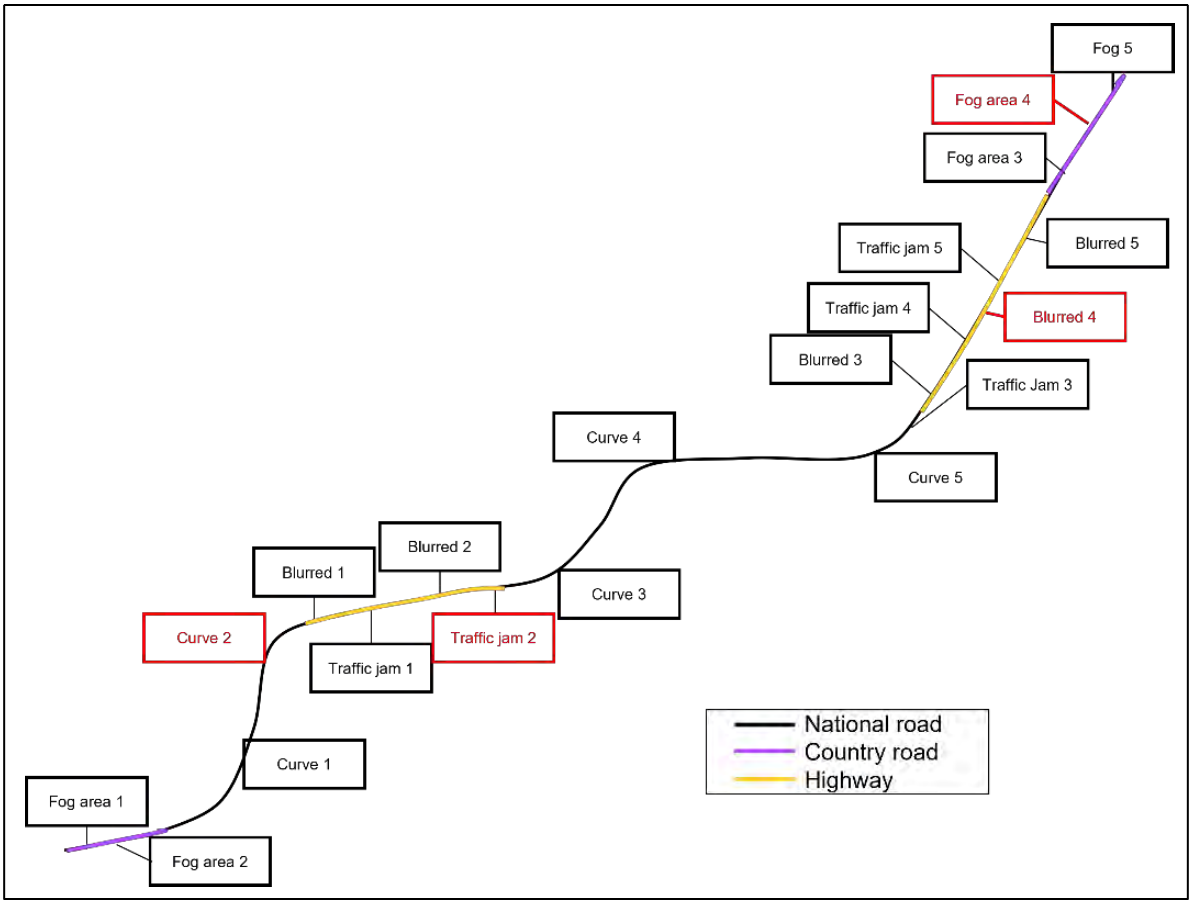
Our choice of eye-tracking technique took into consideration the areas fixated by drivers during the event of the scenarios. This measure allowed to evaluate where the participants looked for information depending on the situation. The SMI Eye Tracking Glasses, a pair of glasses equipped with infrared sensors to monitor eye movements (saccades, fixations and blinks) and a frontal camera to record the field of vision were used. The eye-tracking data were recorded at a sampling frequency of 60 Hz. The glasses were connected to a mobile phone (Samsung Galaxy Note 4) that allowed us to power the glasses, calibrate the gaze measures, display the visual behaviour in real time, and store the video and audio recordings. Eye tracking data were extracted using BeGaze 3.7 software. We also used this software to map the fixations. This mapping consisted in associating each recorded fixation with an AOI and was carried out by a third-party project partner. BeGaze software then calculated the fixation count and duration for each AOI.

2.2.5. Driving Scenario

The driving scenarios were created by Nervtech enterprise for the purpose of this study. They depicted highways, national roads or country roads, with moderate surrounding traffic. Between each type of road, the vehicle was teleported. The new type of road was announced on the screen of the simulator and a black screen preceded the entry to the new type of road. A lead vehicle was always present in front of the driver. The scenarios were created to simulate a realistic road situation, but also to control as much as possible the occurring events and replicate them for all participants. To do so, 20 events occurred, each one separated from each other by 90 seconds. This duration was inspired by Beller et al. (2013). Each event lasted from 10 to 20 seconds, for a total duration of 30 minutes of driving. The events were (1) bent roads, (2) traffic jams, (3) erased road markings and (4) foggy areas (see Figure 28). These events were chosen according to Renault Clio manual because of the possible risk to face a suspension of automation depending on the characteristics of the situation. Among the 20 events, 16 had characteristics that allowed automated systems to function normally, 4 had characteristics that provoked a suspension of automated systems (i.e., one per type of event). These 4 events were randomly placed in the scenario and were identical for all participants. Each type of event was represented an equal number of time (i.e., five times by the type of event).

Figure 28

Representation of the driving scenario with the type of road and the type of event. Events marked in black did not affect automated systems. Events marked in red suspended automated systems.

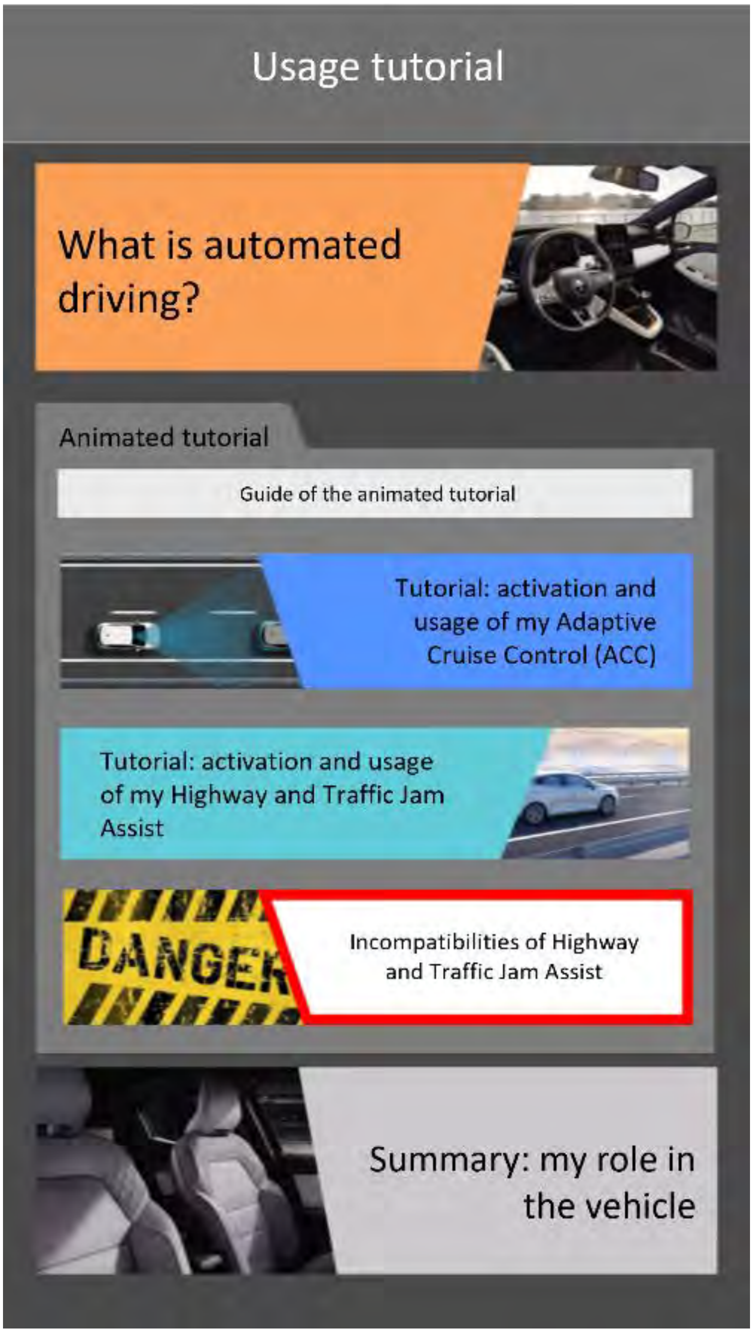


2.2.6. Tutorial

Prior to driving in the simulator, an interactive tutorial realized on Adobe Xd (version 44.1.12.5) was read by the participants. This tutorial aimed to synthesize the functioning of the automated systems. Participants were instructed to read every page of the tutorial for around ten minutes. The tutorial was composed of four parts: (1) an explanation of what an automated system is, (2) the procedure to activate the automated systems, (3) a presentation of limits of the automated systems, and (4) a summary of the tutorial with the experimenter (see Figure 29). (3) The description of the limits of automated systems consisted in presenting the four situations that the driver could encounter during the experiment: bend roads, traffic jams, erased road markings, and areas of fog. For each situation, a description of the course of the event was given step by step, with the presented interfaces, and the actions required to avoid a hazardous situation.

Figure 29

Main page of the tutorial with all chapters that were presented.



2.2.7. Instructions

Participants were instructed to drive as much as possible with their HTJA system activated during a 30-minutes driving scenario. They could deactivate it whenever considered necessary but had to reactivate it as soon as possible if the situation allowed it. They had to follow a white vehicle, maintain constant distance with it and not cross it. Participants were instructed to drive at maximum legal speed.

2.2.8. Familiarization Scenario

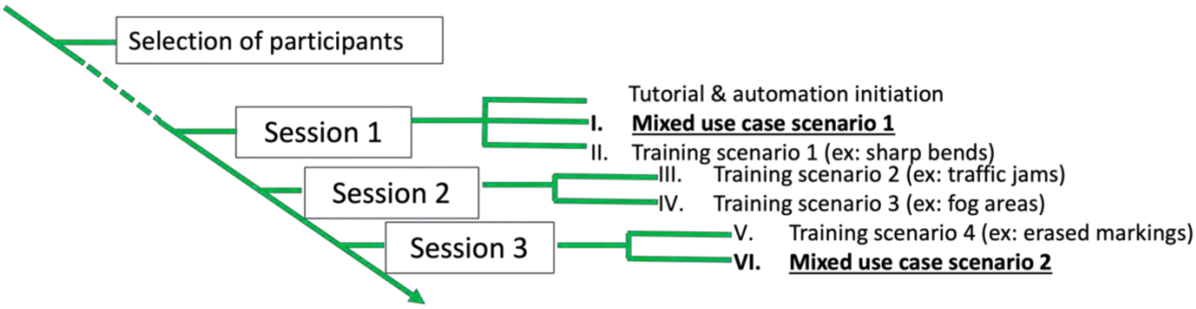
A familiarization scenario was completed by the participants to initiate them to driving the simulator and to automated driving systems. This scenario was a 2x2 straight highway without other vehicles and which lasted around 10 minutes. During this training, the experimenter, who was sited behind the participants, helped participants to get used to the sensations offered by the simulator. They began by turning the steering wheel at low velocity, slowly increasing speed and testing the brakes. Then, the experimenter guided the participants into activating and deactivating automated systems, changing the target speed and front vehicle distance, and informed them about the interfaces that communicated of the state of automated systems. They witnessed the said interfaces in action and were finally confronted to a situation during which automated systems suspended suddenly. They were warned and prepared to act accordingly.

2.3. Procedure

Before recruitment, participants filled out questionnaires regarding their driving habits. If they were selected for the experiment, they were sent explanations regarding the experiment and the functioning of Level-2 automated vehicles. Participants' appointments were fixed for three driving sessions, one per week for three weeks. Each session lasted around 2 hours and 30 minutes, resulting in a total of 7 hours and 30 minutes of experiment per participant. At the beginning of the first session, they filled an informed consent form. Then, the procedure of the experiment was explained. Before driving the simulator, participants were presented the tutorial. They were instructed to read each page of the tutorial at their own pace for around 10 minutes and were free to ask questions. The familiarization scenario was completed and a first questionnaire regarding trust toward automated vehicles was filled. Participants then began the first experimental scenario. Before each experimental driving scenario, the eye-tracking glasses were mounted on the participants and were calibrated. For all participants, the first experimental scenario was a mixed situation scenario (see Figure 30) The experimenter was outside the car and did not intervene except to deal with technical issues. Once the scenario was over, the participants were interviewed. They then filled a mental model questionnaire, workload and trust rating scales. Four scenarios followed and were focused on specific use cases (e.g., scenarios composed only of bend roads). A Latin-square design was used to ensure that all orders of scenarios were completed an equivalent number of repetitions. During the last session, a mixed situation scenario was performed, similar interviews to the previous scenario took place, as well as a semi directed interview regarding opinions of the participants toward the different interfaces. Afterwards, they filled the mental model questionnaire, the workload, and trust rating scales. Participants were thanked and paid for their participation.

Figure 30

Representation of the procedure of the experiment.



2.4. Experiment design

A 2 (between-participants) × 2 (within-participants) experimental design was used. The first factor was the group of interface group and was a between subject factor (VBI vs. MILA). The second factor was the scenario and was a within subject factor (mixed scenario 1 vs. mixed scenario 2).

2.5. Measures

2.5.1. Driving behaviour

The stimulation software allowed to gather vehicle parameters. The measures were similar to Langlois & Soualmi (2016) and represented the driving performances after a period of automated driving. Following Kurpiers et al. (2020) propositions, good driving performances after the suspension of automated systems is an indicator of accurate mode awareness. The mean distance to the center of the lane (i.e., distance between the center of the car and the center of the vehicle in meters) was measured after the suspension of automation in the bent road, fog areas and erased road markings scenarios. TH was measured around five seconds before reaching the speed of a slow vehicle in traffic jams.

2.5.2. Ocular Movements

Gaze positions were coded on a reference image featuring eight AOIs (see Figure 31). For each AOI, the mean fixation duration and number of fixations were extracted during time windows of varying duration depending on the use case (see Table 31). During these time windows, the driver could have perceived the use case on the road, perceive the information on the interface, react to the use case, and return to a nominal situation. The measures during these time widows were divided into five periods: (0) preventive information (for curve only); (1) normal road before the apparition of the event; (2) information of limits of automation for MILA; (3) suspension of automated systems; (4) restoration of a normal situation.

Figure 31

Areas of interest evaluated by the eye-tracking device.

Table 31

Time-windows during which the measures were extracted, depending on the use case.

	Bend road	Traffic jam	Fog area	Erased markings
Before the event	30	5	10	10
After the event	40	35	30	30

The AOIs were gathered into two groups: on-path (on the road) and instrument's cluster. The proportion of fixation duration on-path (on the road) was calculated by dividing the duration of fixation on-path by the total fixation duration. This measure was calculated during the period that followed the suspension of automated systems. The proportion of fixation on the instrument's cluster was calculated by dividing the duration of the fixation on the instrument's cluster by the total duration of fixation. This measure was calculated during periods that preceded the suspension of automated systems.

2.5.3. Rating scales

Mental models

A rating scale was designed to evaluate the mental model of participants regarding the functioning of automated systems depending on the encountered uses cases. It was inspired by the mental model rating scale for Level 2 and Level 3 vehicles of Foster et al. (2019). It was all composed of 11 points Likert-style scales ranging from 0 ("*strongly disagree*") to 10 ("*strongly agree*"). It included 17 items, of which 12 items covered the understanding of automated systems' functioning (3 for each use cases) and 5 items served as distractors. For each item of interest, one end of the rating scale was correct and the other one was incorrect (see Table 32 for detailed items of interest). Mixed linear models were used on each item of interest to evaluate the effect of each variable. The scenario and the interface were used as fixed factors. The participant variable was used as a random factor.

Table 32

Describing the affirmation of the mental model rating scale.

Use case	Affirmation	Correct answer	Type of knowledge
Curve	HTJA is able to function in any type of bend.	strongly disagree	Existence of a limit of automation
	In sharp bends, HTJA becomes unavailable than reactivates itself after the bend.	strongly agree	Presence of auto-activation
	In sharp bends, the ACC turn to suspended state.	strongly disagree	Which automated system suspends
Erased road markings	When road markings are completely or very faded, the ACC and becomes unavailable, then reactivates itself.	strongly disagree	Presence of auto-activation
	When the road markings are completely erased, HTJA will ask you to take over the steering wheel.	strongly agree	Existence of a limit of automation
	When the road markings are removed, the ACC suspends.	strongly disagree	Which automated system suspends
Traffic jams	The ACC is able to brake to match the speed of the vehicle being followed, regardless of the speed of the vehicle being followed.	strongly disagree	Existence of a limit of automation
	The ACC is suspended when braking is too important for the system.	strongly agree	Which automated system suspends
	When an important braking occurs, the ACC is suspended.	strongly agree	Which automated system suspends
Fog areas	The HJTA is able to operate regardless of fog density.	strongly disagree	Existence of a limit of automation
	The ACC is suspended when the fog is too dense, then reactivates itself.	strongly disagree	Presence of auto-activation
	When the fog is too dense, the HJTA suspends.	strongly agree	Existence of a limit of automation

Raw task load index

A French version of the Raw Task Load indeX (RTLX; Cegarra & Morgado, 2009) workload rating scale was completed after the first and last mixed scenarios. This rating scale consists in evaluating the workload in six dimensions (i.e., mental demand, physical demand, temporal demand, effort, performances, and frustration). Participants evaluated the workload of the driving task on a Likert-style rating scale ranging from 0 ("*low*") to 10 ("*high*") for each dimension (see Appendix E for detailed questions). The total workload index was calculated by summing up the ratings of each dimension.

Trust in Automation

Situational trust in automation was evaluated with a personalized question. Participants rated their degree of agreement to the affirmation "*I trusted the Highway and Traffic Jam Assist during this scenario*" on a Likert-style ratings scale ranging from 1 ("*Not at all*") to 10 ("*Totally*"). This rating scale was completed after each driving scenario. The mean ratings of trust in automation were calculated for the mixed scenario 1 and the mixed scenario 2 for each group of interfaces. The progression of trust between the mixed 1 and mixed 2 was assessed by subtracting the ratings of the two scenarios. The difference of progression between the two interfaces' groups was compared with a Mann-Witney's U, normality of residues has not been respected ($p < .05$).

2.5.4. Analysis

The measures of driving performances, ocular behaviour and mental workload were analysed with mixed linear models. Measures of driving performances, and visual fixation, and mental models were analysed separately for each use cases. They will be presented separately for each use case. The following variables were integrated as fixed factors in the model: scenario (mixed scenario 1 vs. mixed scenario 2); interface (VBI vs. MILA). The interactions between these factors were also integrated. The participant factor was integrated as a random factor. Bonferroni's post-hoc tests were carried out when interactions were significant.

3. Results

3.1. Driving Performances After Suspension of Automated Systems

The driving performances were analysed and described separately for each type of use cases (see Table 33 for summary). In the bent road use cases, the mixed linear model analysis revealed a significant effect of the scenario $F(1,39) = 12.56$, $p = 0.001$ on the mean central lane deviation. Participants deviated less from the centre of the lane during mixed scenario 2 (M = 0.46; SD = 0.45) than during mixed scenario 1 (M = 0.63; SD = 0.46). No effect of the interface and no interaction was found significant. During the fog use case, a main effect of the interface was found on the mean central lane deviation $F(1,38) = 8.25$, $p = 0.007$. Participants of the MILA interface ($M = 0.43$; $SD = 0.25$) deviated less than participants of the VBI group ($M = 0.29$; $SD = 0.16$). No effect of the scenario nor interaction effect were found significant ($p > .05$). On roads where lane markings was completely erased, a significant effect of scenario was found $F(1,39) = 10.17$, $p = 0.003$ on the mean central lane deviation. Participants deviated less during the mixed scenario 2 (M = 1.07; SD = 0.56) than during the mixed scenario 1 (M = 1.42; SD = 0.63). The effect of the interface was not significant, neither was the interaction in this use case ($p > .05$). In the traffic jams use cases, no effect of the scenario, of the interface, and interactions were found significant on the TH ($p > .05$).

Table 33

Descriptive statistics (mean (SD)) of the driving performances measures depending on the use case, the scenario, and the interface condition.

Use case (metric)	Scenario	Interface condition	
		MILA	VBI
Bent road (mean central distance)	Mixed scenario 1	1.12 (0.58)	0.81 (0.36)
	Mixed scenario 2	0.64 (0.47)	0.65 (0.39)
Traffic jams (minimal time headway)	Mixed scenario 1	2.61 (0.84)	2.20 (1.33)
	Mixed scenario 2	2.25 (0.57)	2.20 (0.69)
Fog area (mean central distance)	Mixed scenario 1	0.30 (0.19)	0.43 (0.35)
	Mixed scenario 2	0.29 (0.14)	0.43 (0.14)
Erased road markings (mean central distance)	Mixed scenario 1	1.48 (0.69)	1.36 (0.56)
	Mixed scenario 2	1.01 (0.57)	1.14 (0.56)

3.2. Eye-Tracking Measures

3.2.1. Proportion of Fixation on the Instrument's Cluster Prior to Suspensions of Automation

Table 34 details the descriptive statistics of fixation proportion depending on the experimental conditions. In bent road use cases, the mixed linear model revealed a significant effect of the interface $F(1,39) = 11.01$, $p = 0.001$ on the proportion of fixation on the instrument's cluster. Participants with the MILA interface ($M = 0.10$; $SD = 0.15$) looked more at the instrument's cluster before a suspension of automated systems compared to participants of the VBI ($M = 0.02$; $SD = 0.04$). No effect of the scenario and no interaction was revealed for this use case. In the erased marking scenario, the mixed linear model revealed a significant effect of the interface $F(1,39) = 6.09$, $p = 0.018$. Participants with the MILA ($M = 0.13$; $SD = 0.14$) looked more at the instrument's cluster before a suspension of HTJA compared to participants of the VBI ($M = 0.05$; $SD = 0.14$). No effect of the scenario and no interaction was revealed for this use cases. In the traffic jam and fog area use cases, no effect of the interface, of the scenario and interactions were found significant ($p > .05$).

Table 34

Descriptive statistics (mean (SD)) of the proportion of visual fixation on the instrument's cluster before the suspension of automation depending on the use case, the scenario, and the interface condition.

Use case	Scenario	Interface condition	
		MILA	VBI
Bent road	Mixed scenario 1	0.42 (0.36)	0.09 (0.16)
	Mixed scenario 2	0.50 (0.32)	0.08 (0.16)
Traffic jams	Mixed scenario 1	0.29 (0.19)	0.35 (0.22)
	Mixed scenario 2	0.32 (0.25)	0.27 (0.15)
Fog area	Mixed scenario 1	0.08 (0.11)	0.08 (0.15)
	Mixed scenario 2	0.05 (0.08)	0.06 (0.12)
Erased road markings	Mixed scenario 1	0.15 (0.17)	0.06 (0.19)
	Mixed scenario 2	0.10 (0.08)	0.04 (0.06)

3.2.2. Proportion of Fixation On-Path After Suspension of Automation

Table 35 details the descriptive statistics of fixation proportion depending on the experimental conditions. No main effects of the interface and the scenario were found significant for any of the use cases. However, in the erased markings scenario, an interaction effect was revealed by the linear mixed model $F(1, 36.6) = 4.86$, $p = 0.034$. For mixed scenario 2, the difference of fixation proportion between MILA's participants ($M = 0.82$; $SD = 0.12$) and VBI's participants ($M = 0.88$; $SD = 0.06$) seemed to be more important than for mixed scenario 1, for MILA's participants ($M = 0.86$; $SD = 0.09$) and for VBI's participants ($M = 0.83$; $SD = 0.12$). Post-hoc tests revealed no significant differences between each condition ($p > .05$).

Table 35

Descriptive statistics (mean (SD)) of the proportion of visual fixation on the exterior environment after the suspension of automation depending on the use case, the scenario, and the interface condition.

Use case	Scenario	Interface condition	
		MILA	VBI
Bent road	Mixed scenario 1	0.86 (0.15)	0.91 (0.11)
	Mixed scenario 2	0.84 (0.19)	0.88 (0.13)
Traffic jams	Mixed scenario 1	0.79 (0.21)	0.72 (0.15)
	Mixed scenario 2	0.80 (0.20)	0.82 (0.15)
Fog area	Mixed scenario 1	0.79 (0.21)	0.72 (0.15)
	Mixed scenario 2	0.80 (0.20)	0.82 (0.15)
Erased road markings	Mixed scenario 1	0.86 (0.08)	0.83 (0.12)
	Mixed scenario 2	0.82 (0.12)	0.88 (0.06)

3.3. Mental Model Rating Scale

Results of the mixed linear models are reported for the questions regarding each use cases that yield significant effects (see Table 36 for a summary of descriptive statistics). For the bend road use cases and the affirmation "The *HTJA is able to function in any type of bend.",* a significant effect of the interface was observed $F(1,39) = 5.11$, $p = 0.029$. Participants of the MILA group answered better ($M = 6.24$; $SD = 2.98$) than participants of the VBI group ($M =$

4.30; *SD* = 3.50). A main effect of the scenario was also found for this question $F(1, 39) = 5.82$, $p = 0.021$. Participants had a more accurate mental model after the mixed scenario 2 ($M = 5.93$; $SD = 3.58$) than after the mixed scenario 1 ($M = 4.66$; $SD = 3.05$). No interaction was found significant ($p > .05$). For the traffic jams use cases, a main effect of the interface was observed for the question "*The ACC is able to brake to match the speed of the vehicle being followed, regardless of the speed of the vehicle being followed.*" $F(1,39) = 4.93$, $p = 0.032$. Participants of the MILA group had a better mental model ($M = 3.76$; $SD = 3.50$) than the participants of the VBI group ($M = 2.00$; $SD = 2.64$). No effect of the scenario and no interaction were found significant for this question ($p > 0.05$). A significant effect of the scenario was found for the question "*The ACC is suspended when braking is too important for the system*." $F(1,39) = 4.30$, $p = 0.045$. The participants had a better mental model after the mixed scenario 2 ($M = 8.20$; $SD = 2.44$) than after the mixed scenario 1 ($M = 6.98$; $SD = 2.95$). For the fog area use cases, for the affirmation "*The HJTA is able to operate regardless of the fog density.*", a main effect of the scenario was found $F(1,39) = 8.52$, $p = 0.006$. Participants had better mental models in mixed scenario 2 ($M = 7.66$; $SD = 3.16$) than in mixed scenario 1 ($M = 6.17$; $SD = 3.19$). No main effect of the interface and no interaction effect were found for this question (p > .05). For the affirmation "*When the fog is too dense, the HJTA suspends.*", a main effect of the scenario was found $F(1,39) = 5.04$, $p = 0.030$. Participants had a better mental model after mixed scenario 2 ($M = 7.98$; $SD = 3.00$) than after mixed scenario 1 ($M = 6.51$; $SD = 3.21$). For all questions on erased road markings, no significant effect was observed ($p > .05$)

Table 36

Descriptive statistics (mean (SD)) of the scores[1] *to the mental model rating scales depending on the use case, the type of knowledge investigated by the question, the scenario and the interface condition.*

Use case	Type of knowledge	Scenario	Interface condition	
			MILA	VBI
Bent road	Existence of a limit of automation	Mixed scenario 1	5.48 (2.68)	3.80 (3.24)
		Mixed scenario 2	7.00 (3.13)	4.80 (3.75)
		Mixed scenario 1	4.14 (3.05)	5.85 (3.25)

	Which automated system suspends	Mixed scenario 2	5.05 (4.06)	6.00 (4.06)
	Presence of auto-activation	Mixed scenario 1	6.43 (2.80)	5.45 (3.24)
		Mixed scenario 2	6.29 (3.51)	4.65 (3.39)
Traffic jams	Existence of a limit of automation	Mixed scenario 1	3.57 (3.40)	1.25 (1.80)
		Mixed scenario 2	3.59 (3.67)	2.75 (3.14)
	Which automated system suspends	Mixed scenario 1	7.33 (2.52)	6.60 (3.36)
		Mixed scenario 2	7.95 (2.48)	8.45 (2.44)
	Which automated system suspends	Mixed scenario 1	8.24 (2.53)	6.55 (3.66)
		Mixed scenario 2	7.10 (3.27)	7.95 (3.28)
Fog area	Existence of a limit of automation	Mixed scenario 1	6.76 (2.84)	5.55 (3.49)
		Mixed scenario 2	7.57 (3.33)	7.75 (3.06)
	Existence of a limit of automation	Mixed scenario 1	7.10 (3.71)	5.90 (3.71)
		Mixed scenario 2	7.38 (3.68)	8.60 (1.96)
	Presence of auto-activation	Mixed scenario 1	4.71 (3.33)	5.35 (3.57)
		Mixed scenario 2	4.62 (4.18)	6.15 (3.77)
Erased road markings	Existence of a limit of automation	Mixed scenario 1	7.38 (2.69)	7.50 (3.00)
		Mixed scenario 2	8.48 (2.03)	8.45 (2.33)
	Which automated system suspends	Mixed scenario 1	4.81 (3.50)	5.05 (3.59)
		Mixed scenario 2	5.33 (4.20)	5.95 (4.10)
	Presence of auto-activation	Mixed scenario 1	2.86 (3.20)	3.35 (3.05)
		Mixed scenario 2	2.00 (3.15)	3.25 (3.45)

[1] *Scores ranged from 0 to 10. Score close to 10 indicate accurate mental models.*

3.4. Subjective Scales

Regarding trust in automation, Mann-Witney's test revealed a significant effect of the interface on the progression of trust across scenarios $U = 122$, $p = 0.017$, $r = 0.42$. The trust of MILA interface's group progressed more importantly (*Mdn* = 1; *IQR* = 2) than the trust of VBI interface's group (*Mdn* = 0; *IQR* = 1). Regarding mental workload, the linear mixed model revealed a significant effect of the interface on the total score of mental workload $F(1,38) = 7.04$, $p = 0.012$. VBI's participants rated mental workload as lower (M = 16.2; SD = 7.71) than MILA's participants (M = 21.6; SD = 7.65). Self-evaluated mental workload reduced significantly between the mixed scenario 1 ($M = 21.4$; $SD = 7.91$) and the mixed scenario 2 ($M = 16.6$; $SD = 7.69$; $F(1,38) = 19.70$, $p < .001$). No interaction was found ($p > .05$)

4. Discussion

4.1. Study Objectives and Hypotheses

This study aimed to evaluate to what extent a MILA stimulates appropriate level of attention and intervention, induce accurate mode awareness, and appropriate trust in automation, compared to a VBI interface. In a driving simulator, drivers reacted to automated systems' suspensions and had either one of the two interfaces. Their driving performances, ocular behaviours, mental models and self-evaluation of trust and mental workload were gathered. The following hypotheses were tested: (1) a MILA interface stimulates a more appropriate level of attention than a VBI interface which would translate in more gaze fixations on the instrument's cluster before the suspension of automated systems, more important mental workload for MILA than VBI on first usages but a decrease after multiple driving sessions; (2) MILA induces a more accurate mode awareness than VBI which would translate in more precise mental models in shorter periods of time for MILA than for VBI, better driving behaviours when automated systems suspend will be better for MILA than for VBI, more important visual fixations on the road when automation suspends for MILA than for VBI; (3) MILA induces a more important trust in automation than VBI.

4.2. Attention Stimulation

Regarding the stimulation of attention, results of ocular behaviour measurements revealed that the drivers gaze before suspension of automation was influenced by the interface. Participants of the MILA interface' gazed more importantly at the instrument's cluster in the bent road and erased markings use cases. In these situations, the MILA interface appeared to have oriented the attention of the driver to the instrument's cluster. This reveals that the monitoring loop of the drivers was solicited and focused on information regarding the state of the vehicle. This would be an indicator that this interface allows to put back the drivers in the loop and potentially avoid out-of-the loop phenomenon in these situations. These results are coherent with those of Monsaingeon et al. (2021), who found that an indicator of limits of automation was taken into account by drivers in their decision of action. However, these results were not observed in the traffic jam and for area use cases. This can be explained by the fact that the visual resources are heavily exploited in these situations. The fog areas demand to focus vision to direct the vehicle and the traffic jams implies to brake importantly to avoid collision with the followed vehicle. However, MILA interface was more demanding than the VBI interface according to the subjective workload measurement. It appears that orienting the attention of the driver on the instrument's cluster has an attentional cost. Both interfaces were rated as causing very low mental workload. With both interfaces, workload decreased after the second scenario, indicating that the cost of using the automated systems and interfaces decreased. It was expected that workload of the MILA interface's group decreased more importantly that the one of VBI. However, increasing workload is not necessarily detrimental for driving performances. The Malleable Resources Theory (Young & Stanton, 2002) postulates that attentional resources are dependent on the difficulty of a task. When the task is too easy, the level of available resources decreases causing degradation of performances. An interface that induces more important workload might increase the difficulty of the driving task and avoid cognitive underload. Overall, the MILA interface induces an appropriate level of attention in the bent road and erased marking situations.

4.3. Mode Awareness

The interface influenced the behavioural and knowledge dimensions of mode awareness proposed by Kurpiers et al. (2020). The quality of mental models was more important for the MILA interface regarding knowledge about automation's behaviour in bent roads and traffic jams. These results regarding traffic jams are in line with Seppelt and Lee (2019) who found that continuous displays induced more accurate mental models on limits while travelling at slow speed. However, the mental models were not better for MILA's group regarding the erased marking and fog areas. In these situations, information on the instrument's cluster was presented for a short period of time (3 seconds), on the contrary to traffic jam (8 seconds) and bent road (20 seconds). Participants might not have been able to read the information and were not fully aware of the limits and action to perform in these situations. This highlight that time is necessary for the drivers to integrate complex information. Moreover, an area of the instrument's cluster was highlighted to indicate that only the LCA will suspend, and another area indicated that both the ACC and LCA will suspend. Drivers might have not perceived the difference, explaining why questions regarding mental models on which system suspended depending on the situation did not lead to better results. Results also revealed that knowledge regarding limits of automation was influenced by long-term usage of automated systems. These results were coherent with Forster et al. (2019) findings who showed that multiple driving sessions are necessary to form accurate mental models. However, it was expected that mental model's formation would be faster with MILA's group thanks to the indicator of limits of automation, which was not observed. It appears therefore that MILA interface had an immediate effect on mental models and was not influenced by multiple usage.

Regarding driving performances, participants had better driving performances after the suspension of automation in the fog area scenario. This effect was not influenced by learning effect, meaning that it has occurred during the first interaction with automation. However, MILA's interface did not impact driving performances in all other use cases. In opposition to Seppelt and Lee's results (2019), the multimodal interface did not allow to increase TTC in traffic jam situations where emergency braking was necessary. This can be explained by the fact that in Seppelt and Lee's study, the auditory feedback continuously indicated the limit of automation. Our multimodal interface used discrete auditory signal after reaching the limit of automation.

Regarding visual fixation on the exterior environment when suspension of automation occurred, both interface group performed equally. They both directed their attention toward the exterior environment, whatever the interface was. This is an indication that both groups had an accurate mode awareness in these situations, as they were aware that the mode of automation switched to manual driving and required to look at the road. An interesting point is that MILA interface did not distract the drivers and made them look at the instrument's cluster during a crucial moment.

4.4. Relation Between Mental Models and Driving Performances

Overall, the results of mental model questionnaires and driving performances indicate an asymmetry between knowledge about the system's limit and the application of the correct driving behaviour. Mental models were improved by the multimodal interface regarding the limits of automation in bent road and erased markings, but the driving performances were not improved in these situations. Kurpiers et al. (2020) indicate a link between the knowledge and behaviour dimensions of mode awareness. But how the knowledge pillar interacts with the behavioural pillar? Our results suggest that improved knowledge about the system's limit does not automatically lead to improved driving behaviours. According to Rasmussen's SRK model (1983) knowledge, rule and skill based behaviours are supported by different types information. The knowledge behaviours are based on symbols, while skill and rule behaviours are supported by signs. Our indicator of limits of automation used both symbols and signs. Symbols (icons + text) indicated the cause of the limit of automation that would be reached and the action to perform. The sign was a halo with varying size and color to indicate the proximity to the limits. Our results tend to show that the symbols were more used in the bent road and erased marking situations, leading to better mental models. It appears that symbols were prioritized to the detriment of signs, leading to better mental models but not better driving performances. It might be possible that symbols require more attentional resources than sings and that they cannot be processed in parallel. Another argument in favor of this idea is that in the fog areas, driving behaviour was better for the MILA interface group, while the gazes on the instrument's cluster were not more important. This means that drivers did not acquire the symbols proposed on the instrument's cluster and based their behaviour on signs. Signs proposed in the IPLA might that have been perceived in peripheral vision, inducing better driving performances but not feeding mental models.

4.5. Trust in Automation

Trust in automation was influenced by the interface, leading to a more important increase of trust for MILA group compared to VBI group. This result is coherent with previous work on indicators of reliability and limits of automated systems (Beller et al., 2013; Helldin et al., 2013). The relation between mental models and trust in automation was discussed by Seppelt and Lee (2019), who showed that improved mental models lead to increase of trust. Our results confirm that suggestion. Furthermore, it appears that mental models can improve regarding only specific aspects of automated systems' functioning and it will generally impact trust. Indicating limits of automated systems can lead to small effects on mental models, but a general positive effect on trust in automated systems and therefore a more acceptable technology.

4.6. Limitation

Several limitations can be reported in this study. The main limitation is that some factors might have reduced realism in the drivers' experience. A trade-off between ecology of situations and experimental control was necessary. Passing by 20 events at regular pace during a 30 minutes' drive is not a daily occurrence. The events were very controlled to ensure replicability for each participant. The only vehicles present around the drivers were those that were relevant for the event. Therefore, it missed elements like vehicles crossing or pedestrians to make the situations realistic. Moreover, the rhythm to which the event occurred could have made the driving sequences soporific, creating boredom for drivers and reducing their reaction time to suspensions of automated systems. The fact that drivers were equipped with eye-tracking glasses might also have reduced immersion and reduced the realism of their reactions. We oriented toward this type of eye-tracking device to gather high accuracy measures. Eye tracking devices integrated to the vehicle could make the simulation more immersive but are often less accurate. A second limitation concerns the mental model questionnaires. The exact identical questionnaires were given after each scenario. Even though the relevant questions were mixed with distractive questions, the mental model of the participant might have been forged according to this question. Future studies should aim to develop mental model questionnaires that avoid repetitions.

4.7. Conclusion & Future Research

This study offers a novel insight into how interface design can improve the interaction between a human and an automated driving system. Its originality resides in the fact that novice drivers learnt to use automation and that their experience was evaluated with objective and subjective measures. The participants recruited were representative of the population that buys Level-2 vehicles. The results revealed that multimodal interfaces with limit of automation present positive effect such as stimulating adequate level of attention and intervention. It allows to improve mental models and driving behaviours in risky events, globally improving mode awareness and trust in automated systems. This study highlights that improved knowledge about automated systems does not necessarily lead to improve driving behaviours. It appears that indicators of limits of automation should integrate symbols when it aims to improve mental models, and integrate signs when it aims to improve driving performances. The relation between knowledge and behaviours should be further studied to shed better light on their interaction. Even though this study takes into account the learning of drivers with 3 driving sessions separated with one week, some authors suggested that 2 – 3 weeks of daily use is necessary to master the usage of ACC (Weinberger et al., 2001). More time is maybe necessary to master the usage of ACC coupled with LCA. The answer to how mental models' knowledge transfer to driving behaviours might reside in time. Future studies should investigate long-term usage of multimodal interface with limits of automation to evaluate this transfer.

Points clés

- Une interface multimodale indiquant les limites de l'automatisation a un impact positif sur les modèles mentaux des conducteurs.

- Une interface multimodale indiquant les limites de l'automatisation attire l'attention des conducteurs sur le tableau de bord avant la suspension de l'automatisation, ce qui permet de mieux répartir l'attention qu'une interface visuelle classique.

- Les performances de conduite après la suspension de l'automatisation dans les situations de visibilité limitée sont améliorées par l'interface multimodale indiquant les limites de l'automatisation.

- La confiance dans l'automatisation augmente dans le temps de manière plus importante grâce à l'interface multimodale indiquant les limites de l'automatisation qu'à l'interface visuelle classique.

Key points

- A multimodal interface indicating limits of automation positively impact mental models compared to a classical visual interface.

- A multimodal interface indicating limits of automation attracts attention of drivers to the instrument's cluster before the suspension of automation, leading to better attention allocation than a classical visual interface.

- Driving performances after the suspension of automation are improved by the multimodal interface indicating limits of automation in situations where visibility is limited.

- Trust in automation improves more importantly in time thanks to the multimodal interface indicating limit of automation compared to the classical visual interface

GENERAL DISCUSSION

The main objective of this thesis was to evaluate the effect of multimodal interfaces with indications of reliability on the interaction between humans and partially automated vehicles. A central dimension of this thesis is mode awareness, the capacity for drivers to identify the state of automated systems and build mental models of their functioning. The angle through which this goal was approached was to direct the driver's attention to the state of the automated systems through interfaces at the periphery of central vision.

As a first step, a summary of findings of the state-of-art regarding the literature on mode awareness and interfaces' modalities will be presented. The main issues related to interfaces of existing vehicles will be highlighted, along with the interface solutions that were adopted to address these issues. The development and evaluation of a reliability indicator will be summarized. The development and application of a method of evaluation of mode awareness on auditory and haptic interfaces will be summarized. A summary of the effect of a multimodal interface on the development of mental models and attention allocation over time will be presented. As a second step, the contribution of these findings to existing theoretical models of attention distribution and situational awareness will be discussed. Hypotheses regarding the formation of mental models in relation to interfaces display will be presented, followed by the interaction between mental models and trust in automation. As a third step, the main methodological contributions of this work will be discussed, especially the methods of assessment of mode awareness. As a third step, the implications of these results on interface design will be discussed to highlight possible improvements. Finally, the perspectives of research for the interfaces developed here will be presented.

1. Summary and Synthesis of Main Results

In the first part of this manuscript (Chapter 1, Chapter 2, Chapter 3, and Chapter 4), we reviewed the literature on mode awareness and analysed the existing interfaces. A lack of quantitative data on the effect of interfaces' modalities and reliability displays on mode awareness was highlighted by the literature review, despite promising results. The interfaces' characteristics and efficacy of existing vehicles on the interaction with drivers led to establish development perspectives for our own experimental interface. The second part of the manuscript (Chapter 5, Chapter 6, Chapter 7, Chapter 8, and Chapter 9) aimed to design and evaluate the efficiency of new modalities of interfaces to improve mode awareness. The interfaces' objective was to stimulate mode awareness by orienting attention toward the automation's state and their current or future variations. We designed and evaluated 3 interface modalities: a visual one in focal and peripheral vision, an auditory one, and a haptic one. Each one of them was evaluated to ensure that they were useful, that the information was perceived, comprehended, and allowed to project the future variations of automation's state.

The evaluation of interfaces consisted in 6 studies that measured their utility and their effect on mode awareness. The theoretical section of the present manuscript allowed us to first ensure that on-board interfaces are useful to improve the interaction with automated systems (see Chapter 1). In particular, interfaces indicating the reliability of automation are useful, usable (i.e., they induce limited interferences on drivers' behaviours), and, acceptable. Their usefulness lies on the fact that they improve the reactions to automation's suspensions because the drivers' attention is oriented toward the state of automated systems. The presentation of reliability information in peripheral vision is more efficient because the interference with the driving task is lowered. As a result, mode awareness of drivers is improved by the presentation of reliability information. Mode awareness is itself a phenomenon that has received limited interest from automotive research until recently (Chapter 2). Through the meta-analysis on a limited number of studies, we demonstrated that visual interfaces have a significant impact on mode awareness.

Quantitative data in the literature are not sufficient to reach a definitive conclusion on the effect of multimodal interfaces on mode awareness. In the previous studies, researchers have considered mode confusions as a binary phenomenon: either the drivers know or do not know the state of automation. This Manichean approach hides part of the reality of the driver's interaction with the system, leading to possible erroneous interpretations. When only studying the situations in which the drivers identify an automation's state change from the interface, when the suspensions of automation occur, the only conclusion is that the interface fulfills its purpose. Yet, it is also possible that the drivers estimate that automation are suspended when they are not, because of a misinterpretation of the interface indicating a change of automation's state. To address this issue, we proposed to analyse the identification of the state of automation using the Signal Detection Theory. Based on this new approach, we demonstrated that the effect of interfaces on mode awareness can be better quantified. We then have applied this method to the effect of haptic interfaces on detection of suspensions of automation (Chapter 8). The findings of the literature review highlight mode confusions related to existing vehicles. A research strategy was developed to study the effect of interfaces of current vehicles on mode awareness in order to design interfaces that improve it.

At the root of the problematic addressed in this thesis are the issues related to interfaces of existing vehicles. A research strategy was developed and allowed to apprehend the effect of existing interfaces on the understanding of the vehicle's functioning and attention allocation in order to develop our own interfaces (Chapter 3). This strategy began by comparing the effect of two interfaces of existing vehicles that differed in their design approach. Twenty drivers used commercially available partially automated vehicles on-road for 45 minutes (Chapter 4). Results revealed that a multimodal driver-center interface induce better understanding of the vehicle's functioning than a visual vehicle-center interface. It is also visually more demanding and can cause mode confusions. Overall, the conclusions of the state-of-art are that (1) reliability interfaces were lacking in current partially automated vehicles, (2) other sensory modalities than vision can inform on the state of automation, (3) repeated usage of automation and interfaces are necessary to build accurate mental models of automated systems' functioning. These conclusions led to investigate the effect reliability interfaces, auditory interfaces and haptic interfaces on mode awareness.

In the second part of our empirical contribution, we designed new interfaces and evaluated their efficiency to improve mode awareness and attention distribution. The process of design and evaluation followed a rigours methodology. Each interface was designed separately with a specific goal, evaluated with a method built to ensure they increased mode awareness, and improved to cover the issues highlighted in the evaluation process. It is only when all interfaces proved to be efficient that they were gathered into a multimodal interface. This original and rigorous approach allowed to build a multimodal interface in which all elements composing it are optimised to induced mode awareness. This design and evaluation methodology can be exploited in areas outside of the automotive domain and with objectives other than improving mode awareness.

We have deployed an original methodology in the field of mobility to verify the utility of a reliability indicator when several environmental conditions varied (see Chapter 5). The method was based on the presentation of scenarios, placing the participants into situations in which the colour of an indicator varied to indicate the approach of limits of automation. We first showed that drivers judged that the indicator would have influenced their usage of automated systems when environmental conditions were degraded. Then, the methodology allowed us to define personas, based on the profile of response of participants. As we will discuss later, this methodology revealed to be of a real interest in the exploration of the attitude of individuals in the case of interaction with upcoming mobilities.

Once the utility of the reliability interface was verified, we designed and evaluated it (see Chapter 6). The design of this interface, the IPLA, was based on the unprecedented principle of using the instrument's cluster as a support for an interface, aimed to be perceived with peripheral vision. To test its efficiency to help drivers anticipate automation's suspension, a video-based study was deployed. The protocol allowed drivers to choose a driving behaviour depending on the situation and on the information displayed by the IPLA. The results indicated that participants anticipated suspensions of automated systems. The design of the IPLA was improved to encourage efficient attention distribution and avoid inadequate chose of action. Once integrated to the multimodal interface, it allowed to improve mental models on particular knowledge about automation's limits.

New interfaces were designed for the purpose of indicating the variations of state of automated systems. An evaluation methodology was built based on a well-established theoretical model of situational awareness (Endsley, 1995). This methodology aimed to ensure that the developed interfaces were correctly perceived and comprehended by drivers, the first two levels required to build accurate mode awareness. It consisted of three separate experiments that allow new iterations of the interface if the interface does not meet expected results. The first experiment aimed to ensure that the interface was perceived efficiently by using a same/different task. The second experiment evaluated the comprehension of the interface by using a cued-recall task. The third experiment aimed to evaluate the comprehension of the interface while driving by using a cued-recall task during a task as visually demanding as driving. This evaluation method can be exploited for the design of interfaces in other domains where automation is involved, such as aviation. It revealed efficient to help the design of auditory and haptic interfaces.

The lack of experiments in the literature on haptic interfaces and mode awareness led to the exploration of the utility of this interface to indicate transitions of state of automation. An innovative haptic interface in the steering wheel was designed, relying on two different haptic signals (see Chapter 8). The capacity of this interface to orient attention toward the state of automation was evaluated during a simulator study. Forty participants were confronted with suspensions of automation and had to detect them. Signal detection indices were calculated, based on commission and omission errors to quantify the detection suspension of automation. This new method allowed to discriminate the capacities of each of the two haptic signals to indicate the state of automation. The results led to theoretical considerations regarding the existence of a possible haptic resource channel in which the two haptic signals are processed.

Another major interface development relied on the auditory resource channel. Earcons were developed according to design criteria to ensure they would efficiently represent mode transitions (see Chapter 6). The evaluation method mentioned earlier based on the situational awareness model was applied to the earcons. Three experiments took place to ensure the correct perception, comprehension, and interaction with a visual task. Two of these experiments were carried out on a sample of more than 500 participants. The results revealed that the earcons were perceived and comprehended, while causing minimal perturbations effect on a visual task. This perturbation of the visual task, along with design issues, is probably the cause of mode confusions observed on the study with existing vehicles also using earcons. This interface can only stimulate the perception and comprehension levels of mode awareness, not allowing to form projections of future state of automation, limiting its usefulness.

A central aspect of mode awareness is the development of mental models regarding the systems' functioning. During a longitudinal study, lasting three weeks for each of the 40 participants (including six 30-minutes driving sessions), we tested the effect of the multimodal interface gathering all interfaces tested separately (IPLA, haptic and auditory interfaces) on the interaction with a Level-2 vehicle. Each participant followed a pre-experiment formation regarding the functioning of partially automated vehicles. They were later faced with realistic situations in which automation suspended depending on the quality of environmental conditions. To capture the most of drivers' mode awareness and attention allocation, their mental models were rated, their behaviours and visual fixations were measured, their perceived workload and trust in automation were assessed. These measures were compared at the beginning and at the end of the study. Results revealed improvement of mental models of drivers after the last driving session. The takeovers and quantity of visual fixation on the road after the suspension of automation improved with time.

The effect of the multimodal interface was compared with a classical visual interface. The overall effect of the interface was evaluated regarding the attention allocation, mode awareness and trust in automation. The multimodal interface allowed drivers to reallocate their attention before the suspension of automation and led to better control of the vehicle. The three levels of mode awareness were completed, as drivers were able to perceive, comprehend, and project the future state of automated systems. As a side effect of improved mode awareness, trust in automation increased over time for drivers with the multimodal interface. A discussion of the theoretical implications of the effect of multimodal interfaces on mode awareness, attention allocation, mental model formation, and trust in automation will follow. The utility and usability of methods to study and evaluate mode awareness in relation to interface design will then be discussed. The implication of this thesis's results on interface design will then be discussed, followed by perspectives of future works on the interfaces of partially automated vehicles.

2. Theoretical Implications

We will first discuss the theoretical contributions of our results to the first pillars of mode awareness: the awareness of the state of automation according to the Situational Awareness theory. Awareness of the state of automation is dependent on attention allocation. We will then discuss the effect of multimodal interfaces, especially haptic interface and interfaces in peripheral vision, on attention allocation according to the Multiple Resources model. Regarding mental models, the second pillar of mode awareness, we will discuss how they are impacted by reliability indicators, and what factors of partially automated systems make them more difficult to acquire. Finally, the relation between mental models and trust calibration regarding automated systems will be discussed.

2.1. Multimodal Interfaces and Awareness of the State of Automated Systems

A first dimension of mode awareness is the awareness of the state of automation. Mode confusion is defined as the false estimation of the state of automated systems (Baltzer et al., 2017). Previous authors did not differentiate false estimations that automation is suspended and false estimations that automation is activated, considering both as mode confusions (Eom & Lee, 2015). We propose to complete this conception by nuancing these types or mode confusions as commission errors and omission errors. Omission errors can be dangerous, as the driver might fail to takeover control. Commission errors can be equally dangerous because it can lead the drivers to fight against the automated system for the control of the vehicle.

Considering commission and omission errors allow to quantify the detection of states of automation. Signal detection indices can be calculated based on the number of omission and commission errors (Stanislaw & Todorov, 1999). This solution allows to quantify the awareness of the state of automated systems depending on the interface, as it was done in Chapter 2 and Chapter 8 of this work. We strongly believe that this approach can lead to greater advances in the domain human-automation interaction, by directly using the behaviour of the drivers as an indication of omission or commission error.

The multimodal interface tested in this thesis relied on auditory, haptique and interfaces in peripheral vision to improve awareness of the state of automated systems. The awareness of the state of automation was considered as the correct perception and comprehension of information related to automation's states, and the projection of future state of automation depending on the situation (Endsley, 1995). According to the Situational Awareness Theory, interface design influences situational awareness. We will add more nuance to that by stating that the nature of the interface influences which dimension of situational awareness is impacted by interface design. The correct perception and comprehension of auditory signals (see Chapter 7) and haptic signals (see Chapter 8) indicate their utility to improve mode awareness. However, the signals represented here were punctual and were emitted only when a transition of automation's state was occurring. The nature of these signals limits them to the perceptive and comprehension levels of mode awareness, and it prevents them from reaching the projective level.

On the other hand, indicators of reliability, in focal or peripheral vision, rely on continuous information that led to anticipation of transitions of automated systems (Helldin et al., 2013; Kunze et al., 2019). The continuous nature of this information is able to fulfil all dimensions of mode awareness by being perceived and comprehended as an indication of the state of automation and allowing to project future mode transitions. However, the fact that more information is presented in the instrument's cluster can capture the drivers' eye more importantly (see Chapter 4). In order to have an accurate awareness of the automation's state with minimal distraction, the different interface elements that compose the multimodal interface should each have a purpose. Auditory and haptic interfaces should be used to indicate transitions of states of automation. Their purpose should be to alert of suspension. Continuous information of proximity to the limits of automation should be used to allow to project future state of automated systems. Their purpose should be to inform.

2.2. Multimodal Interfaces and Attention Allocation

According to Wickens' (2008) multiple resources theory, demands of a task distributed on multiple resource channels will lead to better performances than if all demands are focused on a single resource channel. Our results confirm this assumption and previous findings on Level-3 vehicles (Zhang et al., 2019). Interestingly, the multiple resource model does not mention a haptic resources channel. Yet, our results and those of previous studies show that the addition of haptic feedback to visual information improve performances compared to visual only interfaces (Petermeijer, 2017; Zhang et al., 2019). In addition to the visual and auditory sensory channels, the addition of a haptic channel would be a relevant addition to the model. It is yet to be determined whether tactile and kinaesthetic interfaces rely on the same resource channel. The addition of tactile and kinaesthetic information through the haptic channel has resulted in a plateau of performance. It therefore appears that both interfaces use the same resource channel.

The IPLA was implemented in a visual interface. The addition of visual information, even though leading to better understanding of the vehicle's functioning (see Chapter 4), can lead to increase workload (Monsaingeon et al., 2019). In the IPLA, texts and pictograms required focal vision while colourful halos could be perceived in peripheral vision. It was expected that, with multiple usage, the drivers would perceive the information of the IPLA with peripheral vision. According to the multiple resources model (2008), focal and peripheral vision exploit different resources. That would have led to lower mental workload compared to if all demands are processed in central vision. However, no usage of peripheral vision was observed in our studies, and workload was judged as more important with the multimodal interface. This suggests that the IPLA attracted the focal vision of drivers, regardless of their multiple exposition to it. An explanation of these results is the usage of very salient colours and shapes for the IPLA. According to the NSEEV model (Wickens, 2015), the appearing an element in peripheral that is sufficiently salient is likely to attract attention. An IPLA that uses very salient visual information would therefore always be perceived with focal vision.

Visual fixations were studied as an indication of attention distribution along our studies. It allowed to measure the attentional demand of a visual interface while looking for a particular information (Chapter 4), as well as measuring mode awareness (Chapter 9). This mode awareness measurement method was proposed by Kurpiers et al. (2020) and is based on the work of Feldhutter et al. (2019) on shifts of modes of automation between partial automation (Level-2) and conditional automation (Level-3). The comparison of gaze duration between

Level-3 and Level-2 modes allowed these authors to estimate the mode awareness. It is based on the fact that drivers can gaze at secondary tasks in Level-3 and must gaze at the road when in Level-2 mode. This proposition cannot apply in vehicles equipped only with Level-2 modes. Visual fixation distribution in Level-2 vehicles should always be located on the exterior environment. We expended these authors proposition by measuring visual fixations during different periods of an event where automation suspended. We found that drivers that allocate attention on the instrument's cluster right before the suspension of automation make attentional resources available for the upcoming event. It allowed to ensure that the interface's information is perceived. Such attentional distribution led to improvement of mental models.

2.3. Difficulty in the Acquisition of Knowledge Related to Partially Automated Driving Systems

In addition to the awareness of the current state of automation, mode awareness also relies on mental models. The IPLA influenced the intention to deactivate automation when detrimental environmental conditions appeared (see Chapter 5), allowed the drivers to anticipate suspensions of automation (see Chapter 6), and improved their mental models (see Chapter 9). It allowed drivers to establish a relation between the state of automation and the environmental conditions, building knowledge about the system's limits. However, different factors come into play to build mental models. A first factor is the gap between the limits of the system and the human limits may have had a role to play in the acquisition of knowledge.

Two types of limits can be differentiated: limits of detection and limits of action of the automated systems. The capacities of detection of lane markings by the cameras in fog areas are more important than the capacities of human eyes. The limits of detection of automation are therefore more difficult to reach than the limits of perception of Humans. On the other hand, the capacities of action of automation are limited by rules that aim not to cause discomfort for drivers. This results in limits, in sharp bends for example, that are more easily reached by automation than by humans (see section Mode Awareness in Chapter 1). Knowledge about limits of detection is probably more accessible than limits of actions, because automated systems have better capacities than humans. Mode awareness therefore appears to be more difficult to build when an important difference can exist between the limits of the system and the limits of humans. Work needs to be done on estimating the difficulty of knowledge to acquire and the corresponding reliability interface. The present work suggests that reliability interfaces using symbols are more adapted for difficult knowledge. Given that such information attracts drivers' gaze, symbols should be presented when large time windows are available before the suspension of automation.

A second factor that influences the difficulty of knowledge acquisition is most likely the complexity of the relation of the systems composing automation. The fact that the two automated systems that compose Level-2 automated vehicles (i.e., ACC and LCA) have different limits of functioning increase complexity. To our knowledge, the mental models of drivers regarding that this kind of specificity of Level-2 vehicles have not been assessed before. It is probable that simplifying the differences of limits between the two systems will lead to accurate mental models faster. A third factor that influences the acquisition of knowledge regarding automation is the time spent using the automated systems. Even though 10 repetitions of exposure to an automated system are sufficient to build an accurate mental model (Beggiato et al., 2015; Forster et al., 2019), it is highly probable that more important repetitions are necessary to grasp all the subtleties of their functioning. As stated by Weinerberger et al. (2001), between two and three weeks of daily usage are necessary for drivers to properly master ACC systems. We can expect this time to be longer with an ACC coupled with an LCA. An interesting approach to study the development of knowledge regarding automation would be to consider the degree of motor movement involved in the interaction.

Some findings suggest that the involvement of the motor system in the interaction with an object plays an important role in the formation of the knowledge about this object (Downing-Doucet & Guérard, 2014; Labeye et al., 2008). It would be interesting to study knowledge formation regarding automation depending on the presentation of abstract information about limits (e.g., IPLA) and the confrontation to the limits with the involvement of a motor response. These two solutions might have their benefits and drawbacks and by complementary depending on the difficulty of the knowledge that must be acquired. Finally, to characterise the difficulty of knowledge acquisition, multiple solutions could be applied. Questionnaires could be filled by naïve participants, asking them to rate, to their opinion, what is the difficulty to learn and apply a statement describing a fact about automated (e.g., "in sharp bent roads, the lane centering assists suspends while the adaptive cruise control remains active"). Another solution would be to make naïve participants learn statements regarding the automated systems functioning, then interrogate them to objectively measure their learning. Once the difficulty of knowledge is established, several interface design should be tested to evaluate the more efficient to acquire complex knowledge.

2.4. Relation Between Mental Models and Trust in Automation

Mental models regarding automation's functioning appear to be bound to trust in automated driving systems. Before usage, trust in automation is a factor that can differentiate drivers in their usage of automation. People sceptical about automation, who deactivate automated systems very often, place less trust in it than people enthusiastic about automated systems (Chapter 5). The representation they have of the functioning of the system leads them not to place their trust in it. When mental model is improved over time thanks to interfaces providing information on automation's limits, trust also increases (Chapter 9). Similarly to previous authors (Seppelt & Lee, 2019), we conclude that better knowledge about the automation's functioning lead to more trust in it. Fortunately, we did not witness over trust in the automation that led to dangerous driving behaviours. We can speculate on the possibility that some over-trusting drivers would rely too much on their interface to estimate when it is time to take-over control of the vehicle because of a hazardous situation. Information about automation that is presented as highly reliabile lead to smaller attention allocation on monitoring the activity of the systems (Avril et al., 2021). Yet, reliability interfaces are not a 100% reliable, because they depend on the capacities of sensors. A solution probably resides in presenting the reliability interfaces during training or tutorials as an informative tool and not a safety one.

3. Methodological Contributions

Multiple original methods have been developed in the process of this thesis to study the effect of the interface on mode awareness. An overview of utility and usability of these methods will be presented. The innovative method of attitude evaluation regarding yet to come interfaces will be discussed first. The method of evaluation of earcons to ensure efficient mode awareness will then be discussed. Will follow a discussion on the qualities and defaults of the mode awareness assessment methods employed in this thesis. Finally, the application of the SDT to mode awareness measurement will be discussed, and improvements will be proposed.

3.1. Scenario Based Method

Several methodological contributions can be noted in this thesis. The first one is the application of a scenario based method, derived from Anderson's (2013) IIT, to the subject of automated driving. This method is based on the usage of attitudes rather than behaviours. Through attitudes, it was possible to directly investigate the persons' representations and not rely on behaviours, which are indirect depictions of peoples' thoughts. This method also enabled to build persona. Thanks to cluster data analysis, profiles of respondents were generated on objective data. This solution offers a great advantage over traditional qualitative creation of persona, sometimes submitted to bias of the experts building them (Mesgari et al., 2019). Finally, this method has a substantial applicative potential, whether it is for the utility interfaces, the comfort in automated vehicles (Delmas et al., 2022) or the price of apples (Hurgobin et al., 2019). Regarding interface design, this method offers great advantages when designing an interface without having the capacity to develop it yet. More generally, it can be applied to technologies, such come automated highly automated vehicles, that do not exist yet. It is highly probable that original studies on new mobilities will see light using this method.

3.2. Mode Awareness Evaluation Method of Icons

The second methodological contribution was the creation of an evaluation method for auditory feedback. This method proposes to objectively evaluate the capacity of earcons to influence mode awareness. It is rooted in Endsley's (1995) model of situational awareness and proposes to assess the first two levels of the mode: the perception and comprehension levels. It is structured in 3 experiments. The first is a same/different task and allows to ensure that the earcons are differentiated. The second is a cued recall task and allows to evaluate that the meaning of the earcons is retrieved. The last experiment is a cued recall task coupled with a dual visual task. If one of these experiments does not show positive results, the earcons need to be redesigned. An advantage of this method is that it can be used on other signals. It was applied in this these on the haptic interface. Moreover, it is not reduced to the automotive domain. It could be applied to aviation, or all other domains than use automation and interfaces to indicate the mode of automation.

3.3. Mode Awareness Measurements

Several measures of mode awareness exist. An objective measure of mode awareness found in Chapter 2, used by multiple authors and that allows to gather quantitative data is based on the freeze prob technique. In a driving simulator, Lee & Ahn (2015) paused the driving situations and asked participants to recall the current mode of automation of their vehicles. A bias of this measure is that it forces the drivers to verbalize and reconstruct thoughts, which does not reflect how drivers would react in real life situations. As discussed previously in this manuscript, mode awareness is a multidimensional construct. A triangulation of measure appears to be necessary to evaluate it as a whole. In the studies presented in this manuscript, several measures were tested: reaction time to questions of automated systems, performances of detection and reaction time to detections of suspensions of automation, eye-tracking measures, and evaluation of the quality of driving behaviours. The reaction times to questions on the automation, asked while driving in Chapter 4, led to interesting findings regarding the accessibility of information on the instrument's cluster. The measure was based on audio recordings. The extraction of data was time consuming and the accuracy was only up to a hundredth of a second. This solution appears to be viable only if the data available are recorded responses to questions.

Another tested measure was the quality of detection of suspensions of automation and the reaction time associated with it. In the simulator study of Chapter 8, participants pressed a button to indicate when they detected that automation suspended. This allows to gather the precise reaction time separating the suspension and the detection. This measurement method reaches its limits when participants have to detect the suspension of multiple automated systems. It would imply for the participant to press different buttons depending on the system that suspended and for them to learn the correct button of each system. A disadvantage of this measure is the reduction in ecology of the situation due to the pressing of the button. Moreover, some situations, such as bent roads, make it more difficult to react appropriately and simultaneously press the button.

Another measurement method applied was eye-tracking measures. In vehicles equipped with Level-3 and Level-2 automation, it is possible to compare the gaze proportion on the outside environment between Level 2 and Level 3 modes. A more important gaze proportion in Level 2 mode than Level 3 mode indicates accurate mode awareness. In vehicles equipped with only Level-2 automation, as it was our case, this statement cannot be applied. To deal with this, we assessed the proposition of gaze on the instrument's cluster and on the exterior environment depending on the moment of the suspension. Drivers that gaze at the instrument's cluster right before the suspension of automation is most likely to back to the controlling loop. If the drivers gaze at the outside environment after the suspension, they most likely understood that the suspension occurred. This measure gathers precise measures but is highly time-consuming.

To complete the eye-tracking measures, the driving performances were assessed after the suspension of automation. This reflected the quality of control the drivers have over the vehicle. If the drivers have an accurate mode awareness and are prepared for the suspension, they will have efficient control over the vehicle. Its addition the driving performances after suspension, we expected to use the quality of actions to apply the Signal Detection Theory.

3.4. Signal Detection Theory Application

Finally, an important methodological contribution of this thesis is the application of the signal detection theory to the assessment of mode awareness in Chapter 8. Usual methods of mode awareness assessment are rating scales, freeze prob techniques, eye-tracking and behaviours (see Chapter 2 for more details). However, these methods often amalgamate omission and commission errors. The signal detection theory allows to consider both types of errors (Janssen et al., 2019). It allows to calculate indices that reflect the capacity of a driver to discriminate the state of automated systems. This method allowed to show that tactile feedback induce better detection of automation's suspension (Chapter 8). It could be applied to any other interface type and any automated systems. The only issue is that in our experiment, the participants had to press a button to express their detection, which is not realistic and intrusive. In the study of Chapter 9, we intended to classify actions of drivers as either appropriated to the situation or not. The goal was to establish correct detection of suspensions, false alarms, misses and correct rejections, making possible to calculate signal detection indices. However, it revealed to be complicated to perform such categorisation because several driving behaviours could be classified as adapted, making impossible to correctly determine the ratio of correct detection and false alarms. More experimental controls are necessary, leaving only one adapted and one unappropriated behaviour for each situation.

4. Implications for Interface Design

The interfaces of reliability in peripheral vision, the auditory interface and the haptic interfaces led to promising results to improve mode awareness. Yet, improvements are possible. We will present possible improvements for the reliability interface first, based on the results on mental models. Then, improvements of the auditory and haptic interface to make them more usable will be presented.

4.1. Improvements of the Reliability Display

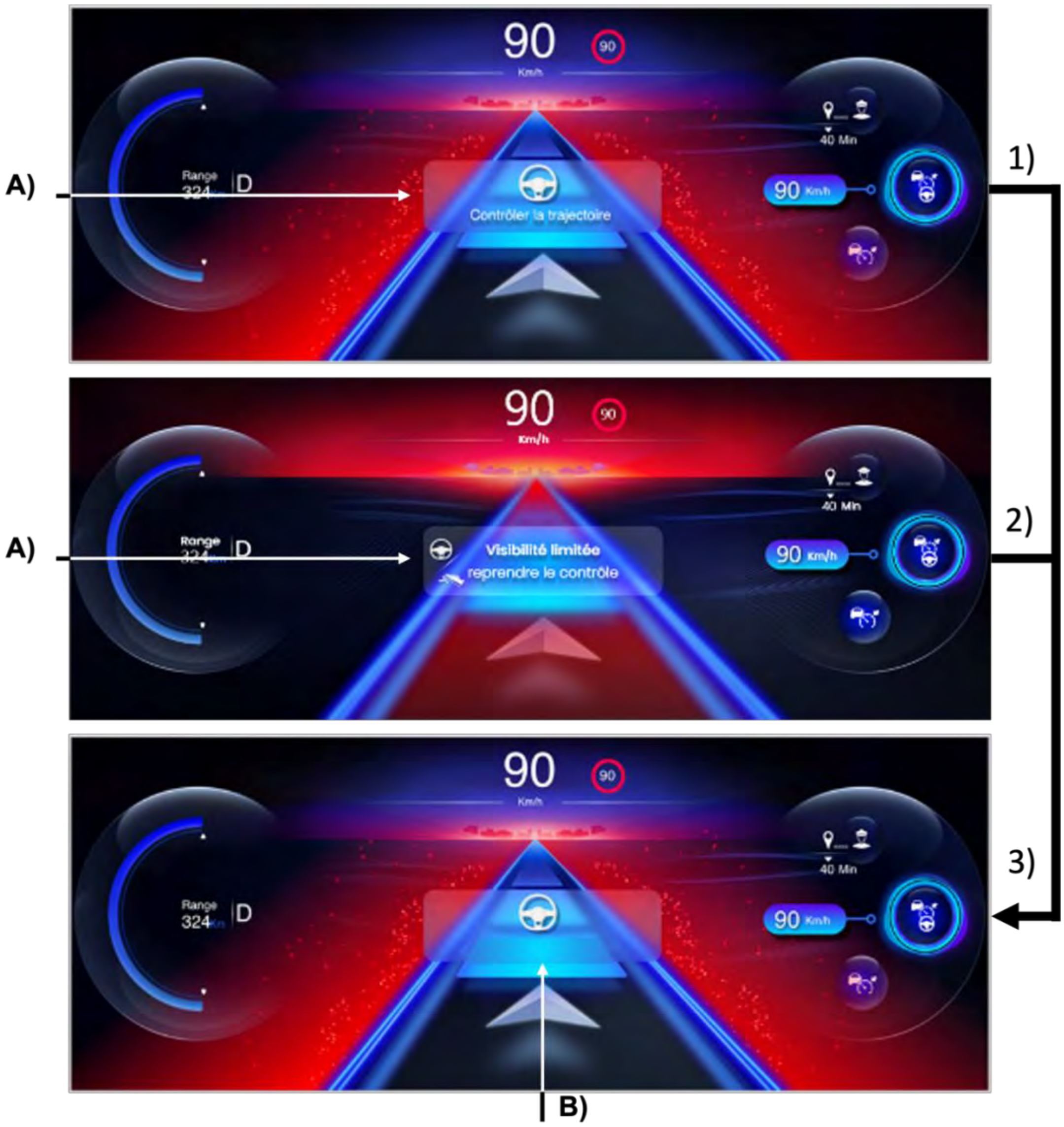
The IPLA developed in this thesis took the form of a halo in the instrument's cluster that evolved in size and colour (i.e., ranging from blue to red) depending on the proximity to the limits of automation. Limits of the LCA and of the ACC were distinguished by displaying the halos at different locations. Two axes were used: the horizontal axis for limits of LCA and vertical axis for the limits of ACC. Information regarding the cause of a limit and actions to perform were displayed at the centre of the instrument's cluster. It was expected that this IPLA required focal vision on first usage and then exploit peripheral vision after prolonged exposition. It was also expected that the IPLA would improve knowledge of drivers regarding which automated systems reach their limits. By judging the eye tracking measures of the same study and the mental model scores, the IPLA was used in focal vision and did not lead to better knowledge about what system reached its limits (i.e., ACC or LCA). To improve the efficiency of the IPLA, we propose to simplify its design. The limits of only the LCA should be displayed, the alerts already present in actual vehicles with ACC appear to be as efficient as the IPLA. Plus, it would reduce possible confusion. To reduce the usage of focal vision and maximize the usage of peripheral vision, minimum signs should be used. It should also induce more Skill and Rule based behaviours (see Figure 32 for a proposition of design).

Figure 32

New proposition of IPLA based on results of this thesis.

1) and 2) represent the interface tested in Chapter 9; 3) is a proposition of a new interface.

A) represents the textual pop up of the IPLA tested in Chapter 9; B) represents the proposition of modification of the pop up.



4.2. Improvement of the Auditory and Haptic Interfaces

The auditory and haptic interfaces revealed to have potential to improve mode awareness in Chapter 7 and Chapter 8's studies. The auditory interface was perceived and comprehended efficiently in experiments where it was alone. The signals therefore appear to be well formulated. When using in the multimodal interface, very few drivers reported hearing them.

To avoid risks of misperceptions and annoyance, they should be used only when actions are required. Differentiating the earcons for LCA only versus LCA and ACC suspension was efficient during the unimodal study, but the difference was not perceived during the multimodal experiment. However, it was not annoying, so use it anyway. The haptic interface was efficiently perceived and comprehended in the Experiments where it was tested alone. The tactile interface should be integrated in future vehicles as it improves the detection of the suspension of LCA. The differentiation of tactile feedback between activation and suspension was not perceived in Chapter 8's experiment. To reduce complexity of the interface, only the feedback of the suspension should be used.

5. Perspectives

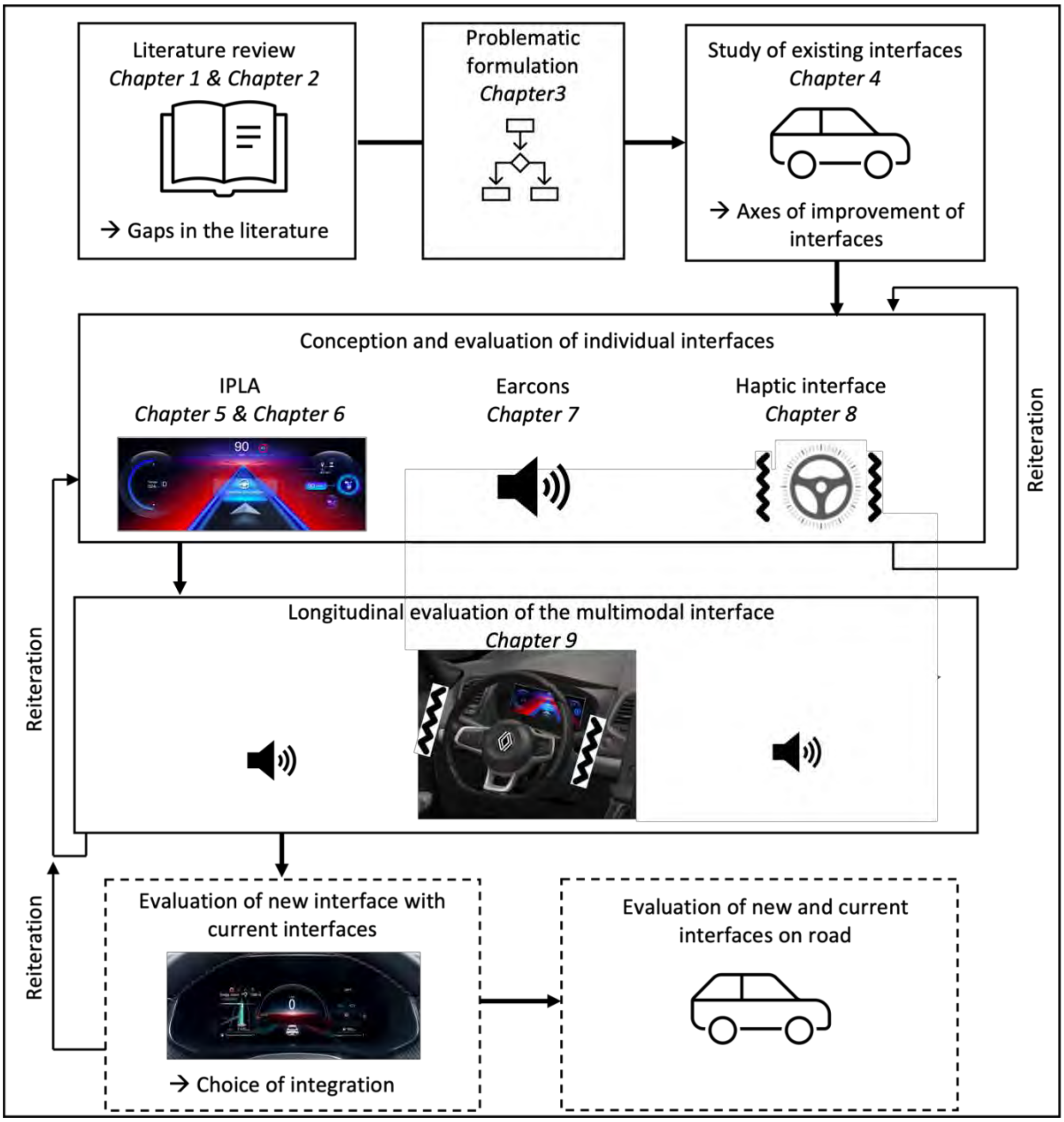
To continue the process of development of interfaces that improve mode awareness in partially automated vehicles, perspectives of integration of the proposed interfaces in the future projects will be presented. Perspectives of evolutions of the IPLA with new interfaces modalities will be presented. The upcoming challenges of integrating the proposed interfaces to existing vehicles will finally be discussed.

The studies carried out in this thesis work followed a step-by-step methodology to ensure that the designed and tested interface respond to the problematic arising from the literature review, and compensate for the issues of existing interfaces. The followed methodology consisted in identifying a problematic in existing literature, studying how interfaces of current vehicles do or do not address this problematic, design and evaluate elements of interfaces that allow to address the problematic, and finally integrate tested elements into one multimodal interface. To pursue this methodology with the aim to integrate the proposed interfaces to vehicles on road, two more steps are necessary. The following step would be to integrate the tested interface to a global interface with all the elements already present in current vehicles. The final step would be to evaluate the integrated interface in real vehicles, on-road (see Figure 33). We will discuss the remaining steps before the final integration in vehicles.

The evaluation of the multimodal interface revealed findings that push toward a reiteration and improvement the IPLA. A version of the IPLA should be tested without those symbols to evaluate its effect on driving behaviours after the suspension of automation. To go further in the direction of using signs and no texts, a LEDs bar at the button of the windshield could duplicate information of the instrument's cluster. It is probable that using a LEDs bar would reveal to be more efficient than using the instrument cluster to transmit information through peripheral vision for two reasons: it would increase the surface area of the display by using all the space of the windshield, and it would place the information closer to the drivers' focal vision. An experiment, using only the signs of the halo and the LEDs bar, with a similar protocol than the one in Chapter 9, should be carried out.

Figure 33

Description of the methodology followed in this thesis to design and evaluate the efficiency of interfaces. Figure 10 has been reused and enhanced with the dotted cells.



Following our work, Renault is pursuing research on the topic of multimodal interfaces. The main concern is now to investigate if the tested interfaces, especially the IPLA, can cohabit with other interfaces present in current vehicles. The main concern is to verify that the proposed interfaces are still efficient when integrated to other interfaces, and whether they are acceptable or not. Regarding visual interfaces, many elements are already present in the instruments cluster, such as the GPS, the Eco-Driving function, the representation of road scene, the state of the speed regulator… It is highly probable that the integration of the IPLA to all this information would increase mental workload. However, it can reveal to be beneficial to drivers' trust and safety. A cost-benefit ratio should be set up to decide if the workload cost of such indicator is worth the benefits. Future studies should be carried out in driving simulator, where the IPLA is integrated to a visual interface as similar as possible to the ones present in current vehicles. The workload should be evaluated, along with trust in automation and driving performances. Then a decision can be made on the integration of such indicator in vehicles. Another element of the interface should be studied before its integration to the other similar interfaces: the haptic interface.

Commercially available vehicles often possess a LDW that takes the form of a high-frequency vibration when a lane marking is crossed. Such information revealed to have a beneficial effect on drivers' reaction to lane departures (Navarro et al., 2016).The tactile feedback investigated in this thesis took the form of a low frequency vibration to indicate the suspension of the LCA. It is plausible that the integration of the new haptic interface to the existing LDW would cause a decrease of performances, as more haptic information is processed at the same time. Future studies should evaluate how the comprehension of the haptic signals is impacted by the presence of both the LDW and the haptic interface of the mode's transition. Laboratory or simulator studies can be carried out to investigate especially the perception and comprehension of haptic signals when they are multiple.

Similar studies can be carried out regarding the auditory signals tested in this thesis. Other auditory signals are present in current vehicles. The perception and comprehension of each auditory interface should be investigated when they are all integrated together. Laboratory studies can be carried out with low financial and human resource cost, using cued recall tasks for each auditory signal, and a procedure similar to the one developed in Chapter 7. Based on results of studies evaluating the interaction between the proposed interface and the existing ones, a choice can be made on their integration. If their integration is judged as useful, they should be tested in a real vehicle on-road. The efficiency of proposed interfaces should then be tested in ecological situations. An important interest should be taken in the acceptability of the interface, as it would determine their final usage by drivers. This step would be the last step in the evaluation of proposed interface and would allow to make a final decision on their integration to vehicles.

CONCLUSION

The purpose of this thesis was to contribute to the understanding of the effect of interface design on the cooperation with automation. In summary, its results showed that interfaces relying on other sensory modality than central vision can improve the cooperation with partially automated vehicles. CMI Project was the support of experimentation for multiple studies on the effect of interfaces on the interaction with automated driving systems. These studies allowed to improve knowledge of Renault company and scientific literature. For Renault, it improved the understanding of the effect of each interface modalities. This work furnishes guidelines to develop interfaces that aim to improve mode awareness in partially automated vehicles. In the long term, these interfaces will contribute to making automation safer and more comfortable for drivers. Moreover, this work provides off-the-shelf methods for future interface evaluation. Regarding the scientific literature, this work contributes to improving knowledge on attentional resources distribution and interface design. It improves knowledge and raises questions about the formation of mental models in relation to interface design and more particularly reliability interfaces.

The results of this thesis, emerging from the collaboration between CLLE Laboratory, a psychological laboratory that works on human factors, and Renault, a major automotive company, illustrate the necessity for this subject to integrate human factors. An important part of this thesis was the collaboration and communication between human factors researchers, engineers, and developers. It was crucial for me, as a researcher in human factors, to adopt the vision and language of engineers that develop the automated systems, in order to work on the integration of such technologies to the everyday drivers. As with most multidisciplinary project, the difficulty resided in the communication between the two different disciplines. Common grounds and definitions must be found. The human factor researchers must understand the functioning of the system they are working on, and the engineers have to understand the necessity of the application of the scientific method to experiments on the users. Fortunately for me, this was facilitated by the open-mindedness of the various members of the project, and by the presence of several human factor researchers, allowing to be heard more easily. On the other hand, it is probable that technologies that emerged during the last decades did not consider the users during their development, leading to poor usability and acceptability.

An example of this is the usability ratings of the software Excel. This software exists for more than three decades, is used by millions of users, and is a very powerful tool. Yet, its usability is judged as just acceptable on the SUS, with an average score of 55 out of 100 (Kortum & Bangor, 2013). It is probable that usability was not tested on the first version of this software, human factor evaluation methods being less applied on tools for the public during the nineties compared to today. Once software is distributed to users, it seems complicated to change the whole structure of it, as they would have to relearn how it operates. It appears that a good start is essential to make tools for humans that are useful and used. I believe that the same can apply to vehicles' inboard technologies. It is best for users to start by using an automated system that can be easily understood and give them a sense of security, to help them create reliable reference points for future use.

Partially automated vehicles are currently used by people who are attracted by the technology or by people who have enough money to buy full option cars. Gradually, automated driving systems should equip the series vehicles, as it is now the case for speed regulator. Even though these technologies currently aim to bring comfort to drivers, some systems have the possibility to improve road safety. For example, the LCA can avoid crashes on the highway by preventing lane departures. Yet, driving with such systems can also imply drowsiness, as the driver does not have much to do except for monitoring the activity of the systems. In a way, it is the snake that bites its own tail: the automated systems avoid risks that they created themselves. One can might ask how these technologies can be useful in a world submitted to global warming, partially because of too important usage of personal vehicles. A brighter future can be imagined thanks to highly automated electric vehicles. Projects like Vilagil in Toulouse was built with the objective to bring new mobilities and durable economic systems in a city that is subjected to pollution and road congestion.

On the long term, highly automated electric vehicles should prevent road accidents, reduce road congestion and pollution. In a completely closed system only relying on automated systems with no exterior parameters interfering, no human error can appear, and accident rate should be close to zero. At one point, the highly automated vehicles will have to interact with other road users. It could raise challenges when no safety drivers are behind the steering wheel of the automated vehicle to avoid hazardous situations or to communicate with other road users. For example, how would pedestrians know if they can cross the road when no eye contact can be made with the driver? External interfaces that communicate to pedestrians probably provide solutions for that kind of challenge. Once again, the human user should be placed at the centre of the research questions on this subject, as automated vehicles are tools at the service of all humans.

If we believe that these new technologies have the possibility to improve the life of users, or even improve society, they should be developed with the objective to be efficient, usable, and acceptable for the user. The goal is always the same, provide information for everyday users and allow them to interact, or even cooperate with the tool. For our subject, we took the part of considering the interface as a window through which the driver glimpses inside the automated systems. Future research should be led with the same idea in mind: a clear window helps the users comprehend and communicate with their tools.

REFERENCES

Anderson, N. H. (2013). Unified Psychology Based on Three Laws of Information Integration. *Review of General Psychology*, *17*(2), 125-132. https://doi.org/10.1037/a0032921

Avril, E., Valéry, B., Navarro, J., Wioland, L., & Cegarra, J. (2021). Effect of Imperfect Information and Action Automation on Attentional Allocation. *International Journal of Human–Computer Interaction*, *37*(11), 1063-1073. https://doi.org/10.1080/10447318.2020.1870817

Baddeley, A. D. (1986). *Working memory*. Oxford University Press.

Bagheri, N., & Jamieson, G. A. (2004). *The impact of context-related reliability on automation failure detection and scanning behaviour*. *1*, 212-217. https://doi.org/10.1109/ICSMC.2004.1398299

Bainbridge, L. (1983). Ironies of automation. In G. Johannsen & J. E. Rijnsdorp (Éds.), *Analysis, Design and Evaluation of Man–Machine Systems* (p. 129-135). Pergamon. https://doi.org/10.1016/B978-0-08-029348-6.50026-9

Baltzer, M. C. A., Rudolph, C., López, D., & Flemisch, F. (2017). Cooperative Guidance and Control in Highly Automated Convoys—StrAsRob. In N. A. Stanton, S. Landry, G. Di Bucchianico, & A. Vallicelli (Éds.), *Advances in Human Aspects of Transportation* (Vol. 484, p. 547-559). Springer International Publishing. https://doi.org/10.1007/978-3-319-41682-3_46

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics*, *68*, 138-145. https://doi.org/10.1016/j.apergo.2017.11.010

Bazilinskyy, P., & de Winter, J. (2015). Auditory interfaces in automated driving : An international survey. *PeerJ Computer Science*, *1*, e13. https://doi.org/10.7717/peerj-cs.13

BEA. (1992). *Rapport de la commission d'enquête sur l'accident survenu le 20 janvier 1992 près du Mont Sainte-Odile (Bas Rhin) a l'Airbus A 320 immatricule F-GGED exploite par la compagnie Air Inter*. https://cir.nii.ac.jp/crid/1571698599239963136

Beattie, D., Baillie, L., & Halvey, M. (2015). A comparison of artificial driving sounds for automated vehicles. *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 451-462. https://doi.org/10.1145/2750858.2807519

Beggiato, M., Pereira, M., Petzoldt, T., & Krems, J. (2015). Learning and development of trust, acceptance and the mental model of ACC. A longitudinal on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *35*, 75-84. https://doi.org/10.1016/j.trf.2015.10.005

Belderbos, C. (2015). *Authority Transition Interface—A Human Machine Interface for Taking over Control from a Highly Automated Truck*. Industrial Design Engineering.

Beller, J., Heesen, M., & Vollrath, M. (2013). Improving the driver–automation interaction : An approach using automation uncertainty. *Human Factors*, *55*(6), 1130-1141.

Bisantz, A. M., & Seong, Y. (2001). Assessment of operator trust in and utilization of automated decision-aids under different framing conditions. *International Journal of Industrial Ergonomics*, *28*(2), 85-97. https://doi.org/10.1016/S0169-8141(01)00015-4

Blömacher, K., Nöcker, G., & Huff, M. (2020). The evolution of mental models in relation to initial information while driving automated. *Transportation Research Part F: Traffic Psychology and Behaviour*, *68*, 198-217. https://doi.org/10.1016/j.trf.2019.11.003

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to Meta-Analysis*. John Wiley & Sons.

Brewster, S. A., Capriotti, A., & Hall, C. V. (1998). Using compound earcons to represent hierarchies. *HCI Letters*, *1*(1), 6-8.

Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1993). *An evaluation of earcons for use in auditory human-computer interfaces*. 222-227. https://doi.org/10.1145/169059.169179

Brooks, J. O., Crisler, M. C., Klein, N., Goodenough, R., Beeco, R. W., Guirl, C., Tyler, P. J., Hilpert, A., Miller, Y., Grygier, J., Burroughs, B., Martin, A., Ray, R., Palmer, C., & Beck, C. (2011). Speed choice and driving performance in simulated foggy conditions. *Accident Analysis & Prevention*, *43*(3), 698-705. https://doi.org/10.1016/j.aap.2010.10.014

Caroux, L., Consel, C., Dupuy, L., & Sauzéon, H. (2018). Towards context-aware assistive applications for aging in place via real-life-proof activity detection. *Journal of Ambient Intelligence and Smart Environments*, *10*(6), 445-459. https://doi.org/10.3233/AIS-180505

Carpenter, S. K., Pashler, H., & Vul, E. (2006). What types of learning are enhanced by a cued recall test? *Psychonomic Bulletin & Review*, *13*(5), 826-830. https://doi.org/10.3758/BF03194004

Carsten, O., & Martens, M. H. (2019). How can humans understand their automated cars? HMI principles, problems and solutions. *Cognition, Technology & Work*, *21*(1), 3-20. https://doi.org/10.1007/s10111-018-0484-0

Cegarra, J., & Morgado, N. (2009). *Étude des propriétés de la version francophone du NASA-TLX*. 233-239.

Christoffersen, K., Hunter, C. N., & Vicente, K. J. (1996). A longitudinal study of the effects of ecological interface design on skill acquisition. *Human Factors*, *38*(3), 523-541. https://doi.org/10.1518/001872096778701917

Cochran, W. G. (1954). The Combination of Estimates from Different Experiments. *Biometrics*, *10*(1), 101-129. JSTOR. https://doi.org/10.2307/3001666

Craig, A. (1979). Nonparametric Measures of Sensory Efficiency for Sustained Monitoring Tasks. *Human Factors*, *21*(1), 69-77. https://doi.org/10.1177/001872087902100109

Cramer, S., & Klohr, J. (2019). Announcing Automated Lane Changes : Active Vehicle Roll Motions as Feedback for the Driver. *International Journal of Human–Computer Interaction*, *35*(11), 980-995.

Cramer, S., Miller, B., Siedersberger, K.-H., & Bengler, K. (2017). Perceive the difference : Vehicle pitch motions as feedback for the driver. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1699-1704. https://doi.org/10.1109/SMC.2017.8122860

Davey, J., Turner, R. M., Clarke, M. J., & Higgins, J. P. (2011). Characteristics of meta-analyses and their component studies in the Cochrane Database of Systematic Reviews : A cross-sectional, descriptive analysis. *BMC Medical Research Methodology*, *11*(1), 160. https://doi.org/10.1186/1471-2288-11-160

Davis, K., Owusu, E. B., Marcenaro, L., Feijs, L., Regazzoni, C., & Hu, J. (2017). Effects of ambient lighting displays on peripheral activity awareness. *IEEE Access*, *5*, 9318-9335. https://doi.org/10.1109/ACCESS.2017.2703866

Dehais, F., Causse, M., Vachon, F., Régis, N., Menant, E., & Tremblay, S. (2014). Failure to detect critical auditory alerts in the cockpit : Evidence for inattentional deafness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(4), 631-644. https://doi.org/10.1177/0018720813510735

Delmas, M., Camps, V., & Lemercier, C. (2022). Effects of environmental, vehicle and human factors on comfort in partially automated driving : A scenario-based study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *86*, 392-401. https://doi.org/10.1016/j.trf.2022.03.012

de Winter, J. C. F., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness : A review of the empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196-217. https://doi.org/10.1016/j.trf.2014.06.016

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys : The tailored design method*. John Wiley & Sons.

Dingus, T. A., Neale, V. L., Klauer, S. G., Petersen, A. D., & Carroll, R. J. (2006). The development of a naturalistic data collection system to perform critical incident analysis : An investigation of safety and fatigue issues in long-haul trucking. *Accident Analysis & Prevention*, *38*(6), 1127-1136. https://doi.org/10.1016/j.aap.2006.05.001

Downing-Doucet, F., & Guérard, K. (2014). A motor similarity effect in object memory. *Psychonomic Bulletin & Review*, *21*(4), 1033-1040. https://doi.org/10.3758/s13423-013-0570-5

Endsley, M. R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32-64. https://doi.org/10.1518/001872095779049543

Endsley, M. R. (2017). Autonomous Driving Systems : A Preliminary Naturalistic Study of the Tesla Model S. *Journal of Cognitive Engineering and Decision Making*, *11*(3), 225-238. https://doi.org/10.1177/1555343417695197

Endsley, M. R., & Garland, D. J. (2000). *Pilot Situation Awareness Training in General Aviation*. *44*, 357-360. https://doi.org/10.1177/154193120004401107

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(2), 381-394. https://doi.org/10.1518/001872095779064555

Eom, H., & Lee, S. (2015). Human-Automation Interaction Design for Adaptive Cruise Control Systems of Ground Vehicles. *Sensors*, *15*(6), 13916-13944. https://doi.org/10.3390/s150613916

Faltaous, S., Baumann, M., Schneegass, S., & Chuang, L. L. (2018). *Design guidelines for reliability communication in autonomous vehicles*. 258-267. https://doi.org/10.1145/3239060.3239072

Farell, B. (1985). « Same »-"Different" Judgments : A Review of Current Controversies in Perceptual Comparisons. *Psychological Bulletin*, *98*(3), 419-456. https://doi.org/10.1037//0033-2909.98.3.419

Feldhütter, A., Härtwig, N., Kurpiers, C., Hernandez, J. M., & Bengler, K. (2019). Effect on Mode Awareness When Changing from Conditionally to Partially Automated Driving. In S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, & Y. Fujita (Éds.), *Advances in Intelligent Systems and Computing* (Vol. 823, p. 314-324). Springer. https://doi.org/10.1007/978-3-319-96074-6_34

Feldhütter, A., Segler, C., & Bengler, K. (2018). Does Shifting Between Conditionally and Partially Automated Driving Lead to a Loss of Mode Awareness? In N. A. Stanton (Éd.), *Advances in Human Aspects of Transportation* (Vol. 597, p. 730-741). Springer International Publishing. https://doi.org/10.1007/978-3-319-60441-1_70

Flemisch, F., Adams, C., Conway, S., Goodrich, K., Palmer, M., & Schutte, P. (2003). *The H-Metaphor as a Guideline for Vehicle Automation and Interaction*. NASA Technical Reports.

Flemisch, F. O., Bengler, K., Bubb, H., Winner, H., & Bruder, R. (2014). Towards cooperative guidance and control of highly automated vehicles : H-Mode and Conduct-by-Wire. *Ergonomics*, *57*(3), 343-360. https://doi.org/10.1080/00140139.2013.869355

Forster, Y., Hergeth, S., Naujoks, F., Beggiato, M., Krems, J. F., & Keinath, A. (2019). *Learning and development of mental models during interactions with driving automation : A simulator study*. *10*, 398-404. https://ir.uiowa.edu/drivingassessment/2019/papers/59

Forster, Y., Naujoks, F., & Neukum, A. (2016). Your Turn or My Turn? : Design of a Human-Machine Interface for Conditional Automation. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - Automotive'UI 16*, 253-260. https://doi.org/10.1145/3003715.3005463

Fruchart, E., & Mullet, E. (2012). Factors involved in the intent to move from one team to another among handball players. *Psicológica*, *33*(3), 425-434.

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., Griffith, L., Oremus, M., Raina, P., Ismaila, A., Santaguida, P., Lau, J., & Trikalinos, T. A. (2011). Conducting quantitative synthesis when comparing medical interventions : AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology*, *64*(11), 1187-1197. https://doi.org/10.1016/j.jclinepi.2010.08.010

Furukawa, H., Inagaki, T., Shiraishi, Y., & Watanabe, T. (2003). Mode awareness of a dual-mode adaptive cruise control system. *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems, Man and Cybernetics. Conference Theme - System Security and Assurance (Cat. No.03CH37483)*, *1*, 832-837. https://doi.org/10.1109/ICSMC.2003.1243918

Gaffary, Y., & Lécuyer, A. (2018). The use of haptic and tactile information in the car to improve driving safety : A review of current technologies. *Frontiers in ICT*, *5*, 5. https://doi.org/10.3389/fict.2018.00005

Gang, N., Sibi, S., Michon, R., Mok, B., Chafe, C., & Ju, W. (2018). Don't Be Alarmed : Sonifying Autonomous Vehicle Perception to Increase Situation Awareness. *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 237-246. https://doi.org/10.1145/3239060.3265636

Gaspar, J., & Carney, C. (2019). The effect of partial automation on driver attention : A naturalistic driving study. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *61*(8), 1261-1276. https://doi.org/10.1177/0018720819836310

Helldin, T., Falkman, G., Riveiro, M., & Davidsson, S. (2013). *Presenting system uncertainty in automotive UIs for supporting trust calibration in autonomous driving*. 210-217. https://doi.org/10.1145/2516540.2516554

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Ho, C., Spence, C., & Spence, C. (2017). *The Multisensory Driver : Implications for Ergonomic Car Interface Design*. CRC Press. https://doi.org/10.1201/9781315555423

Hoc, J. M., & Amalberti, R. (2007). Cognitive Control Dynamics for Reaching a Satisficing Performance in Complex Dynamic Situations. *Journal of Cognitive Engineering and Decision Making*, *1*(1), 22-55. https://doi.org/10.1177/155534340700100102

Hofmans, J., & Mullet, E. (2013). Towards unveiling individual differences in different stages of information processing : A clustering-based approach. *Quality & Quantity*, *47*(1), 455-464. https://doi.org/10.1007/s11135-011-9529-7

Horiguchi, Y., Fukuju, R., & Sawaragi, T. (2006). An Estimation Method of Possible Mode Confusion in Human Work with Automated Control Systems. *2006 SICE-ICASE International Joint Conference*, 943-948. https://doi.org/10.1109/SICE.2006.315649

Howell, D. C. (2012). *Statistical Methods for Psychology*. Cengage Learning.

Huang, Y., & Ferreira, F. (2020). The Application of Signal Detection Theory to Acceptability Judgments. *Frontiers in Psychology*, *11*, 11.

Hurgobin, Y., Le Floch, V., & Lemercier, C. (2019). Effect of multiple extrinsic cues on consumers' willingness to buy apples : A scenario-based study. *Food Quality and Preference*, *81*. https://doi.org/10.1016/j.foodqual.2019.103860

Jacob, R. J. K., & Karn, K. S. (2003). Eye Tracking in Human-Computer Interaction and Usability Research : Ready to Deliver the Promises. In J. Hyönä, R. Radach, & H. Deubel (Éds.), *The Mind's Eye, cognitive and applied aspects of eye movement research* (p. 573-605). North Holland. https://doi.org/10.1016/B978-044451020-4/50031-1

Janssen, C. P., Boyle, L. N., Kun, A. L., Ju, W., & Chuang, L. L. (2019). A hidden markov framework to capture human–machine interaction in automated vehicles. *International Journal of Human–Computer Interaction*, *35*(11), 947-955.

Janssen, C. P., Donker, S. F., Brumby, D. P., & Kun, A. L. (2019). History and future of human-automation interaction. *International Journal of Human-Computer Studies*, *131*, 99-107. https://doi.org/10.1016/j.ijhcs.2019.05.006

Jeon, M. (2019). *Multimodal Displays for Take-over in Level 3 Automated Vehicles while Playing a Game*. 1-6. https://doi.org/10.1145/3290607.3313056

Jeon, M., Davison, B. K., Nees, M. A., Wilson, J., & Walker, B. N. (2009). *Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies*. 91-98. https://doi.org/10.1145/1620509.1620528

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Katzourakis, D. I., de Winter, J. C. F., Alirezaei, M., Corno, M., & Happee, R. (2014). Road-Departure Prevention in an Emergency Obstacle Avoidance Situation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *44*(5), 621-629. https://doi.org/10.1109/TSMC.2013.2263129

Kodinariya, T. M., & Makwana, R. (2013). *Review on determining number of Cluster in K-Means Clustering*. *1*(6), 90-95.

Kortum, P. T., & Bangor, A. (2013). Usability Ratings for Everyday Products Measured With the System Usability Scale. *International Journal of Human-Computer Interaction*, *29*(2), 67-76. https://doi.org/10.1080/10447318.2012.681221

Kramer, G. (1994). *Auditory Display : Sonification, Audification, And Auditory Interfaces*. *19*(2). https://doi.org/10.2307/3680606

Kunze, A., Summerskill, S. J., Marshall, R., & Filtness, A. J. (2019). *Conveying Uncertainties Using Peripheral Awareness Displays in the Context of Automated Driving*. 329-341. https://doi.org/10.1145/3342197.3344537

Kurpiers, C., Biebl, B., Mejia Hernandez, J., & Raisch, F. (2020). Mode awareness and automated driving—What is it and how can it be measured? *Information*, *11*(5), 277. https://doi.org/10.3390/info11050277

Labeye, E., Oker, A., Badard, G., & Versace, R. (2008). Activation and integration of motor components in a short-term priming paradigm. *Acta Psychologica*, *129*(1), 108-111. https://doi.org/10.1016/j.actpsy.2008.04.010

Langlois, S., & Soualmi, B. (2016). *Augmented reality versus classical HUD to take over from automated driving : An aid to smooth reactions and to anticipate maneuvers*. 1571-1578. https://doi.org/10.1109/ITSC.2016.7795767

Lee, J. D., & See, K. A. (2004). Trust in Automation : Designing for Appropriate Reliance. *Human Factors*, *46*(1), 50-80. https://doi.org/10.1518/hfes.46.1.50_30392

Lee, S. H., & Ahn, D. R. (2015). Design and verification of driver interfaces for adaptive cruise control systems. *Journal of Mechanical Science and Technology*, *29*(6), 2451-2460. https://doi.org/10.1007/s12206-015-0536-9

Lemercier, C., Pêcher, C., Berthié, G., Valéry, B., Vidal, V., Paubel, P.-V., Cour, M., Fort, A., Galéra, C., Gabaude, C., Lagarde, E., & Maury, B. (2014). Inattention behind the wheel : How factual internal thoughts impact attentional control while driving. *Safety Science*, *62*, 279-285. https://doi.org/10.1016/j.ssci.2013.08.011

Lemmens, P. M. C., Bussemakers, M. P., & de Haan, A. (2000, avril). *The effect of earcons on reaction times and error-rates in a dual-task vs. A single-task experiment*. 6th International Conference on Auditory Display, Atlanta, GA, United States. http://hdl.handle.net/1853/50685

Lu, Z., & de Winter, J. C. F. (2015). A Review and Framework of Control Authority Transitions in Automated Driving. *Procedia Manufacturing*, *3*, 2510-2517. https://doi.org/10.1016/j.promfg.2015.07.513

Mankoff, J., Dey, A. K., Hsieh, G., Kientz, J., Lederer, S., & Ames, M. (2003). *Heuristic evaluation of ambient displays*. 169-176. https://doi.org/10.1145/642611.642642

Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, *18*(1), 50-60. https://doi.org/10.1214/aoms/1177730491

McNicol, D. (1972). *A primer of signal detection theory*. Allen and Unwin.

Mesgari, M., Okoli, C., & de Guinea, A. O. (2019). Creating rich and representative personas by discovering affordances. *IEEE Transactions on Software Engineering*, *45*(10), 967-983. https://doi.org/10.1109/TSE.2018.2826537

Meyer, J. (2004). Conceptual Issues in the Study of Dynamic Hazard Warnings. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(2), 196-204. https://doi.org/10.1518/hfes.46.2.196.37335

Michon, J. A. (1985). A critical view of driver behavior models : What do we know, what should we do? In L. Evans & R. C. Schwing (Éds.), *Human Behavior and Traffic Safety* (p. 485-524). Springer US. https://doi.org/10.1007/978-1-4613-2173-6_19

Mirnig, A. G., Gärtner, M., Laminger, A., Meschtscherjakov, A., Trösterer, S., Tscheligi, M., McCall, R., & McGee, F. (2017). Control Transition Interfaces in Semiautonomous Vehicles : A Categorization Framework and Literature Analysis. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '17*, 209-220. https://doi.org/10.1145/3122986.3123014

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses : The PRISMA Statement. *PLOS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Monk, A. (1986). Mode errors : A user-centred analysis and some preventative measures using keying-contingent sound. *International Journal of Man-Machine Studies*, *24*(4), 313-327. https://doi.org/10.1016/S0020-7373(86)80049-9

Monsaingeon, N., Carli, Y., Caroux, L., Langlois, S., & Lemercier, C. (2021, juillet). Indicating the limits of partially automated vehicles with drivers' peripheral vision. *12th International Conference on Applied Human Factors and Ergonomics - AHFE 2021*. https://hal.archives-ouvertes.fr/hal-03114194

Monsaingeon, N., Caroux, L., Langlois, S., & Lemercier, C. (2021, juin). Auditory feedbacks for automated driving : An evaluation method. *21st Triennial Congress of International Ergonomics Association - IEA 2021*. https://hal.archives-ouvertes.fr/hal-03114244

Monsaingeon, N., Caroux, L., Mouginé, A., Langlois, S., & Lemercier, C. (2021). Impact of interface design on drivers' behavior in partially automated cars : An on-road study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *81*, 508-521. https://doi.org/10.1016/j.trf.2021.06.019

Monsaingeon, N., Mouginé, A., Caroux, L., Langlois, S., & Lemercier, C. (2019, juillet). Étude comparative de véhicules semi-autonomes en conduite réelle : Effet du tableau de bord sur les émotions et la charge mentale du conducteur. *Colloque de psychologie ergonomique EPIQUE*. https://hal-univ-tlse2.archives-ouvertes.fr/hal-02094110

Murata, A., & Kuroda, T. (2015). Effects of auditory and tactile warning on drivers' response to hazard under noisy environment. In D. Harris (Éd.), *Engineering Psychology and Cognitive Ergonomics* (p. 45-53). Springer International Publishing. https://doi.org/10.1007/978-3-319-20373-7_5

Nadri, C., Ko, S., Diggs, C., Winters, M., Sreehari, V. K., & Jeon, M. (2021). Novel Auditory Displays in Highly Automated Vehicles : Sonification Improves Driver Situation Awareness, Perceived Workload, and Overall Experience. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *65*(1), 586-590. https://doi.org/10.1177/1071181321651071

Naujoks, F., Forster, Y., Wiedemann, K., & Neukum, A. (2017). A Human-Machine Interface for Cooperative Highly Automated Driving. In N. A. Stanton, S. Landry, G. Di Bucchianico, & A. Vallicelli (Éds.), *Advances in Human Aspects of Transportation* (Vol. 484, p. 585-595). Springer International Publishing. https://doi.org/10.1007/978-3-319-41682-3_49

Navarro, J., Yousfi, E., Deniel, J., Jallais, C., Bueno, M., & Fort, A. (2016). The impact of false warnings on partial and full lane departure warnings effectiveness and acceptance in car driving. *Ergonomics*, *59*(12), 1553-1564. https://doi.org/10.1080/00140139.2016.1158323

Noah, B. E., Gable, T. M., Chen, S.-Y., Singh, S., & Walker, B. N. (2017). Development and preliminary evaluation of reliability displays for automated lane keeping. *9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '17*, 202-208. https://doi.org/10.1145/3122986.3123007

Noah, B. E., & Walker, B. N. (2017). Trust calibration through reliability displays in automated vehicles. *ACM/IEEE International Conference on Human-Robot Interaction - HRI '17*, 361-362. https://doi.org/10.1145/3029798.3034802

Onimaru, S., & Kitazaki, M. (2010). Visual and tactile information to improve drivers' performance. *2010 IEEE Virtual Reality Conference (VR)*, 295-296. https://doi.org/10.1109/VR.2010.5444759

Parasuraman, R., & Riley, V. (1997). Humans and automation : Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230-253. https://doi.org/10.1518/001872097778543886

Payre, W., Cestac, J., Dang, N.-T., Vienne, F., & Delhomme, P. (2017). Impact of training and in-vehicle task performance on manual control recovery in an automated car. *Transportation Research Part F: Traffic Psychology and Behaviour*, *46*, 216-227. https://doi.org/10.1016/j.trf.2017.02.001

Petermeijer, S., Doubek, F., & de Winter, J. (2017). Driver response times to auditory, visual, and tactile take-over requests : A simulator study with 101 participants. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1505-1510. https://doi.org/10.1109/SMC.2017.8122827

Petermeijer, S. M. (2017). *A vibrotactile interface to support the driver during the take-over process* [PhD Thesis]. Technische Universität München.

Politis, I., Brewster, S. A., & Pollick, F. (2014). Evaluating multimodal driver displays under varying situational urgency. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 4067-4076. https://doi.org/10.1145/2556288.2556988

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*(3), 257-266. https://doi.org/10.1109/TSMC.1983.6313160

Regan, M. A., Lee, J. D., & Young, K. L. (Éds.). (2009). *Driver distraction : Theory, effects, and mitigation*. CRC Press.

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ*, *342*. https://doi.org/10.1136/bmj.d549

Rothstein, H. R., & Bushman, B. J. (2012). Publication bias in psychological science : Comment on Ferguson and Brannick (2012). *Psychological Methods*, *17*(1), 129-136. https://doi.org/10.1037/a0027128

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 5-19. https://doi.org/10.1518/001872095779049516

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. *Handbook of Human Factors and Ergonomics*, *2*, 25.

Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Medical Informatics and Decision Making*, *7*(1), 16. https://doi.org/10.1186/1472-6947-7-16

Scott, J. J., & Gray, R. (2008). A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Human Factors*, *50*(2), 264-275. https://doi.org/10.1518/001872008X250674

Seppelt, B. D., & Lee, J. D. (2007). Making adaptive cruise control (ACC) limits visible. *International Journal of Human-Computer Studies*, *65*(3), 192-205. https://doi.org/10.1016/j.ijhcs.2006.10.001

Seppelt, B. D., & Lee, J. D. (2019). Keeping the driver in the loop : Dynamic feedback to support appropriate use of imperfect vehicle control automation. *International Journal of Human-Computer Studies*, *125*, 66-80. https://doi.org/10.1016/j.ijhcs.2018.12.009

Seppelt, B., & Victor, T. (2020). Driver's mental model of vehicle automation. In D. L. Fisher, W. J. Horrey, J. D. Lee, & M. A. Regan (Éds.), *Handbook of Human Factors for Automated, Connected, and Intelligent Vehicles* (p. 55-66). CRC Press.

Serna, R. W., Preston, M. A., & Thompson, G. B. (2013). Assessing nonverbal same/different judgments of auditory stimuli in individuals with intellectual disabilities : A methodological investigation. *Revista Brasileira de Analise Do Comportamento*, *5*(2), 19.

Sivak, M. (1996). The Information That Drivers Use : Is it Indeed 90% Visual? *Perception*, *25*(9), 1081-1089. https://doi.org/10.1068/p251081

Solís-Marcos, I., Ahlström, C., & Kircher, K. (2018). Performance of an additional task during level 2 automated driving : An on-road study comparing drivers with and without experience with partial automation. *Human Factors*, *60*(6), 778-792. https://doi.org/10.1177/0018720818773636

Solís-Marcos, I., & Kircher, K. (2019). Event-related potentials as indices of mental workload while using an in-vehicle information system. *Cognition, Technology & Work*, *21*(1), 55-67. https://doi.org/10.1007/s10111-018-0485-z

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149. https://doi.org/10.3758/BF03207704

Strand, N., Stave, C., & Ihlström, J. (2018). *A case-study on drivers' mental model of partial driving automation*. 25th ITS World Congress.

Suzuki, K., & Jansson, H. (2003). An analysis of driver's steering behaviour during auditory or haptic warnings for the designing of lane departure warning system. *JSAE Review*, *24*(1), 65-70. https://doi.org/10.1016/S0389-4304(02)00247-3

Tardieu, J., Misdariis, N., Langlois, S., Gaillard, P., & Lemercier, C. (2015). Sonification of in-vehicle interface reduces gaze movements under dual-task condition. *Applied Ergonomics*, *50*, 41-49. https://doi.org/10.1016/j.apergo.2015.02.004

Terken, J., Riener, A., Schroeter, R., & Osswald, S. (2013). *Adjunct Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*.

Trémolière, B., & Djeriouat, H. (2016). The sadistic trait predicts minimization of intention and causal responsibility in moral judgment. *Cognition*, *146*, 158-171. https://doi.org/10.1016/j.cognition.2015.09.014

Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., Lee, J. D., & Markkula, G. (2015). *Analysis of Naturalistic Driving Study Data : Safer Glances, Driver Inattention, and Crash Risk*. Transportation Research Board. https://trid.trb.org/view/1323585

Vidulich, M. A., & Tsang, P. S. (2015). The confluence of situation awareness and mental workload for adaptable human–machine systems. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 95-97. https://doi.org/10.1177/1555343414554805

Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics*, *30*(3), 261-293. https://doi.org/10.3102/10769986030003261

Wandtner, B. (2018). *Non-driving related tasks in highly automated driving-Effects of task characteristics and drivers' self-regulation on take-over performance*.

Wang, J., & Soffker, D. (2019). Bridging Gaps Among Human, Assisted, and Automated Driving With DVIs : A Conceptional Experimental Study. *IEEE Transactions on Intelligent Transportation Systems*, *20*(6), 2096-2108. https://doi.org/10.1109/TITS.2018.2858179

Wang, J.-S., Knipling, R. R., & Goodman, M. J. (1996). The role of driver inattention in crashes : New statistics from the 1995 Crashworthiness Data System. [Data set]. In *40th Annual Proceedings of the Association for the Advancement of Automotive Medicine*. American Psychological Association. https://doi.org/10.1037/e533322008-001

Weinberger, M., Winner, H., & Bubb, H. (2001). Adaptive cruise control field operational test—The learning phase. *JSAE Review*, *22*(4), 487-494. https://doi.org/10.1016/S0389-4304(01)00142-4

Wickens, C. D. (2008). Multiple Resources and Mental Workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 449-455. https://doi.org/10.1518/001872008X288394

Wickens, C. D. (2015). Noticing events in the visual workplace : The SEEV and NSEEV models. In R. R. Hoffman, P. A. Hancock, M. W. Scerbo, R. Parasuraman, & J. L. Szalma (Éds.), *The Cambridge Handbook of Applied Perception Research* (p. 749-768). Cambridge University Press. https://doi.org/10.1017/CBO9780511973017.046

Wickens, C. D., Gempler, K., & Morphew, M. E. (2000). Workload and Reliability of Predictor Displays in Aircraft Traffic Avoidance. *Transportation Human Factors*, *2*(2), 99-126. https://doi.org/10.1207/STHF0202_01

Wille, M., & Debus, G. (2005). Regulation of speed and time-headway in traffic. In G. Underwood (Éd.), *Traffic & transport psychology : Theory and Application.* (p. 327-337). Elsevier.

Wintersberger, P., Dmitrenko, D., Schartmüller, C., Frison, A.-K., Maggioni, E., Obrist, M., & Riener, A. (2019). S(C)ENTINEL : Monitoring automated vehicles with olfactory reliability displays. *Proceedings of the 24th International Conference on Intelligent User Interfaces - IUI '19*, 538-546. https://doi.org/10.1145/3301275.3302332

Young, M. S., & Stanton, N. A. (2002). Malleable attentional resources theory : A new explanation for the effects of mental underload on performance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *44*(3), 365-375. https://doi.org/10.1518/0018720024497709

Zhang, B., de Winter, J., Varotto, S., Happee, R., & Martens, M. (2019). Determinants of take-over time from automated driving : A meta-analysis of 129 studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 285-307. https://doi.org/10.1016/j.trf.2019.04.020

APPENDICES

Appendix A: Chapter 4 – On road questions

Questions cible
Quelle est la vitesse actuelle ? (autopilot actif)
Quelle est la vitesse actuelle ? (autopilot inactif)
Est-ce que l'AutoPilot est actif ? (autopilot actif)
Est-ce que l'AutoPilot est actif ? (autopilot inactive)
Est-ce que le régulateur adaptatif est actif ?
Est-ce que le centrage voie est actif ?
Questions personnelles
Avez-vous des enfants/petits enfants ?
Avez-vous des animaux ?
Quels sont leurs noms ?
Faites-vous du sport ?
Questions intérieures
Quelle est la couleur du stylo avec lequel j'écris ?
Quel est la radio que nous écoutons ?
Quelle est la couleur de mes vêtements ?
Questions extérieures
Pensez-vous qu'il va pleuvoir aujourd'hui ?
Est que le véhicule à notre droite est vert ?
Est-ce que le véhicule derrière nous est noir ?

Appendix B: Chapter 5 – Sample of stories presented to participants

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très clair**. La route est en **ligne droite**. Les lignes blanches au sol **sont bien marquées**. Le voyant PLS est **vert**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très clair**. La route est en **ligne droite**. Les lignes blanches au sol **sont bien marquées**. Le voyant PLS est **orange**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très clair**. La route est en **ligne droite**. Les lignes blanches au sol **ne sont pas bien marquées**. Le voyant PLS est **vert**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très clair**. La route est en **ligne droite**. Les lignes blanches au sol **ne sont pas bien marquées**. Le voyant PLS est **orange**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très clair**. La route est en **virage serré**. Les lignes blanches au sol **sont bien marquées**. Le voyant PLS est **vert**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très clair**. La route est en **virage serré**. Les lignes blanches au sol **sont bien marquées**. Le voyant PLS est **orange**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très clair**. La route est en **virage serré**. Les lignes blanches au sol **ne sont pas bien marquées**. Le voyant PLS est **vert**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très clair**. La route est en **virage serré**. Les lignes blanches au sol **ne sont pas bien marquées**. Le voyant PLS est **orange**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très pluvieux**. La route est en **ligne droite**. Les lignes blanches au sol **ne sont pas bien marquées**. Le voyant PLS est **vert**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très pluvieux**. La route est en **ligne droite**. Les lignes blanches au sol **ne sont pas bien marquées**. Le voyant PLS est **orange**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très pluvieux**. La route est en **virage serré**. Les lignes blanches au sol **sont bien marquées**. Le voyant PLS est **vert**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très pluvieux**. La route est en **virage serré**. Les lignes blanches au sol **sont bien marquées**. Le voyant PLS est **orange**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très pluvieux**. La route est en **virage serré**. Les lignes blanches au sol **ne sont pas bien marquées**. Le voyant PLS est **vert**.

Julie est sur la route des vacances. Son véhicule conduit sur une route départementale. Le temps est **très pluvieux**. La route est en **virage serré**. Les lignes blanches au sol **ne sont pas bien marquées**. Le voyant PLS est **orange**.

Appendix C: Chapter 6 – Links to the videos used in the experiment

IPLA interface

Scenario	State of automated systems	URL
Bent road	Suspend	Link
Bent road	Stays activated	Link
Traffic jam	Suspend	Link
Traffic jam	Stays activated	Link
Fog	Suspend	Link
Fog	Stays activated	Link
Erased markings	Suspend	Link
Erased markings	Stays activated	Link

Reference interface

Scenario	State of automated systems	URL
Bent road	Suspend	Link
Bent road	Stays activated	Link
Traffic jam	Suspend	Link
Traffic jam	Stays activated	Link
Fog	Suspend	Link
Fog	Stays activated	Link
Erased markings	Suspend	Link
Erased markings	Stays activated	Link

Appendix D: Chapter 7 – Links to the earcons

Earcon label	Automated mode transition	Notes	URL
L2	From Level 0 to Level 2	C2 – C4	Link
L1	From Level 2 to Level 1	C4 – C3	Link
L0	From Level 2 to Level 0	C4 – C2 – C2	Link

Appendix E: Chapter 9 – Raw Task Load indeX

Ce questionnaire est composé de 6 échelles de notation destinées à évaluer votre charge de travail investie durant la réalisation de ces scénarios de conduite. Pour chacune des échelles, veuillez s'il vous plait, marquer d'une croix (X) le niveau qui correspond le mieux à ce que vous avez ressenti.

Exigence mentale
Quelle a été l'importance de l'activité mentale et intellectuelle requise (ex. réflexion, décision, calcul, mémorisation, observation, recherche etc.) ? La tâche vous a-t-elle paru simple, nécessitant peu d'attention (faible) ou complexe, nécessitant beaucoup d'attention (élevée) ?

Faible Elevée

Précisez votre réponse. Quel élément a causé une exigence mentale plus ou moins importante ?

____

Exigence physique
Quelle a été l'importance de l'activité physique requise (ex. pousser, porter, tourner, marcher, activer, etc.) ? La tâche vous a-t-elle paru facile, peu fatigante, calme (faible) ou pénible, fatigante, active (élevée) ?

Faible Elevée

Exigence temporelle
Quelle a été l'importance de la pression temporelle causée par la rapidité nécessitée pour l'accomplissement de la tâche ? Etait-ce un rythme lent et tranquille (faible) ou rapide et précipité (élevé) ?

Faible Elevée

Performance
Quelle réussite pensez-vous avoir eu dans l'accomplissement de votre tâche ? Comment pensez-vous avoir atteint les objectifs déterminés par la tâche ?

Bonne Mauvaise

Effort

Quel degré d'effort avez-vous dû fournir pour exécuter la tâche demandée, (mentalement et physiquement) ?

Faible Elevée

Frustration

Pendant l'exécution du travail vous êtes-vous senti satisfait, relaxé, sûr de vous (niveau de frustration faible), ou plutôt découragé, irrité, stressé, sans assurance (niveau de frustration élevé) ?

Faible Elevée