Running title: PERCEIVED USABILITY OF AUTOMATED VEHICLES

The Utility of Psychological Measures in Evaluating Perceived Usability of Automated Vehicle Interfaces – A Study with Older Adults

Alexandra Voinescu[a,c], Phillip L. Morgan[b], Chris Alford[c], Praminda Caleb-Solly[c]

[a] Alexandra Voinescu, Department of Psychology, University of Bath, University of Bath, Claverton Down, Bath, BA2 7AY, UK. E-mail address: a.voinescu@bath.ac.uk

[b] Phillip L. Morgan, Cardiff University, Human Factors Excellence (HuFEx) Research Group, and Centre for Artificial Intelligence, Robotics, and Human-Machine Systems (IROHMS), School of Psychology, Cardiff, CF10 3AT, UK. E-mail address: morganphil@cardiff.ac.uk

[c] Chris Alford, Psychological Sciences Research Group, University of the West of England, Bristol BS16 1QY, UK. E-mail: chris.alford@uwe.ac.uk

[c] Praminda Caleb-Solly, Bristol Robotics Laboratory, University of the West of England, Bristol BS16 1QY, UK. E-mail: praminda.caleb-solly@uwe.ac.uk

Correspondence concerning this article should be addressed to: Phillip L. Morgan, Cardiff University, School of Psychology, Cardiff, CF10 3AT, UK. E-mail address: morganphil@cardiff.ac.uk

Author Disclosure Statement

No competing financial interests exist.

The Utility of Psychological Measures and Situational Factors in Evaluating Perceived Usability of Automated Vehicle Interfaces – A Study with Older Adults

Abstract

The design of the traditional vehicle human-machine interfaces (HMIs) is undergoing major change as we move towards fully connected and automated vehicles (CAVs). Given the diversity of user requirements and preferences, it is vital for designers to gain a deeper understanding of any underlying factors that could impact usability. The current study employs a range of carefully selected psychological measures to investigate the relationship with self-report usability of an in-CAV HMI integrated into a fully automated Level 5 simulator, during simulated journeys. Twenty-five older adults (65-years+) participated and were exposed to four journeys in a virtual reality fully automated CAV simulator (with video recorded journeys) into which our HMI was integrated. Participants completed a range of scales and questionnaires, as well as computerized cognitive tests. Key measures were: perceived usability of the HMI, cognitive performance, personality, attitudes towards computers, trust in technology, simulator sickness, presence and emotion. HMI perceived usability correlated positively with cognitive performance (e.g., working memory) and some individual characteristics such as trust in technology and negatively with neuroticism anxiety. Simulator sickness was associated negatively with CAV HMI perceived usability. Positive emotions correlated positively with reported usability across all four journeys, while negative emotions were negatively associated with usability only in the case of the last two journeys. Increased sense of presence in the virtual CAV simulator was not associated with usability. Implications for design are critically discussed. Our research is highly relevant in the design of high-fully automated vehicle HMIs, particularly for older adults, and in informing

policy-makers and automated mobility providers of how to improve older people's uptake of this technology.

Keywords: usability, connected automated vehicles, fully automated driverless cars, human machine interface, older adults, individual differences.

The Utility of Psychological Measures and Situational Factors in Evaluating Perceived Usability of Automated Vehicle Interfaces – A Study with Older Adults

1. Introduction

Automated vehicles[1] (AVs) and Connected and automated vehicles (CAVs)[2,3] are developing at a fast pace, with car manufactures and technology companies such as Google, Uber, Tesla, General Motors, Toyota, Hyundai, Audi, BMW and Volvo rushing to try and get them deployed on roads (Mounce & Nelson, 2019). According to the Society of Automotive Engineers (SAE, 2016), there are five levels of automation ranging from 1 to 5, with level 5 being fully automated. In terms of highly and fully AVs, driving-related actions are performed by the automated driving system under most (Level 4) or all (Level 5) roadway and environmental conditions (Riener, Boll, & Kun, 2016; SAE, 2016; Talebpour & Mahmassani, 2016). Some expect that by 2040, AVs will reach around 25% of the global new-vehicle market (MIT Reviews, 2017) and UK regulations anticipate that

[1] A vehicle which is capable of fulfilling the operational functions of a traditional car without a human operator (SAE, 2016)

[2] A vehicle which can communicate with other vehicles and infrastructure systems (e.g., Vehicle-to-Vehicle and Vehicle-to-Infrastructure communications) (Talebpour & Mahmassani, 2016).

[3] Connected vehicle technology can be applied and is independent of various levels of automation, including Level 3 and fully automated Level 5 vehicles, and is proposed to be safer and more reliable (Talebpour & Mahmassani, 2016).

vehicles with AV and CAV features will be on public roads by 2021 (BBC News, 2017), though these predictions do not consider the various levels of automation. The use of CAVs has the potential to reduce road traffic accidents, congestion and time in traffic (Fagnant & Kockelman, 2015; Ferati, Murano, & Giannoumis, 2017; Mounce & Nelson, 2019).

Older adults might be one population sector that would likely benefit most from CAVs as they are one of the most vulnerable populations with regards to traffic accidents (Dotzauer, De Waard, Caljouw, Pöhler, & Brouwer, 2015), are more likely to drive less than, for example, younger adults (Choi & Ji, 2015), and more likely to cease driving due to safety and/or enjoyment reasons (Siren & Haustein, 2015). Special attention should be given to the mobility of this population sector (e.g., Morgan et al., 2017) as the percentage of older adults within many countries is increasing at a significant rate (WHO, 2015), and forecasts predict that older adults will continue working in later life which is likely to shape the current transport landscape (Shergold, Lyons, & Hubers, 2015). In response to that, CAVs potentially offer older adults better mobility options for e.g., social, domestic, pleasure and work purposes and avoid social exclusion (Harvey, Guo, & Edwards, 2019; Li, Blythe, Guo, & Namdeo, 2019; Nikitas, Avineri, & Parkhurst, 2018), especially those vehicles that are highly (SAE Level 4) or fully (SAE Level 5) automated. It is estimated that older adults will have the greatest increase in annual vehicle miles traveled with the use of AVs compared to younger population (Harper, Hendrickson, Mangones, & Samaras, 2016). However, some older adults are more concerned particularly in relation to technology failure and giving-up control (especially those that are currently driving, Musselwhite, 2019) and less favorable towards AVs than younger people (Hudson, Orviska, & Hunady, 2019; Hulse, Xie, & Galea, 2018). Musselwhite (2019) found that older adults who gave up driving were more positive about AVs and stressed the importance of maintaining their mobility and connectivity with others and leisure activities. Thus, other studies failed to identify age effects on engagement with Level 4 AV technology (Molnar et al., 2018). Due to divergent attitudes of older adults towards AVs, it is crucial that best practices in CAV human machine interface (HMI) design for older adults (e.g., (Brewer, Garcia, Schwaba, Gergle, & Piper, 2016; Fisk, Czaja, Rogers, Charness, & Sharit, 2009; Orphanides & Nam, 2017) are developed and tested with older adults themselves (Li et al., 2019; Morgan, Voinescu, Alford, & Caleb-Solly, 2018; Morgan et al., 2017) in order to increase acceptance, usability and improve attitudes towards AVs.

1.1.Current Study

The current study is part of a large project funded by the UK Government[4] with major objectives concerning understanding of expectations that older people might have for CAVs and usability issues associated with their use. Our study involves an HMI for a fully automated CAV (Level 5) simulator whose design is based upon best principles for older adults derived from the literature. Part of the novelty of the study stands in its implementation, as most of the previous studies that survey the users' opinions about AVs/CAVs have not actually been designed to give participants experience of highly/fully automated driving situations. Exposing participants to simulated CAV journeys might have different results than those from studies where participants have not experienced simulated journeys (e.g. survey-based studies) and had limited information about how it actually feels to be driven by an AV/CAV (Nordhoff, van Arem, & Happee, 2016). Most importantly, our study is among the first that investigates CAV HMI usability in a virtual reality

[4] XXXX project is a multi-sector collaboration which is helping to advance the successful implementation of Connected and Automated Vehicles (CAVs) in the UK, by developing services and capabilities that link user needs and system requirements. XXXX project seeks to develop products and services that maximise the benefits of Connected and Autonomous vehicles for users and transport authorities. By adopting a user-centred approach, XXXX aims to achieve a better understanding of consumer demands and expectations, including the implications and challenges of an ageing society with a focus on CAV user experience of older adults.

(VR) simulator which adds value to our results as it enhances the ecological validity of the usability evaluation. Usability studies benefit from an ecologically valid evaluation setting, but as with any study that uses a VR setting there are other variables that are likely to be associated with usability outcomes such as presence in VR, simulator sickness and participants' mood and emotions.

The objective of the current study was to investigate the magnitude and direction of relationships between perceived HMI usability, individual differences and situational factors related to the context of usability evaluation. We investigated the relationship between the perceived usability of a fully automated CAV HMI and individual differences such as: cognitive abilities, personality, attitudes towards computers and trust in technology and situational variables like: simulator sickness, presence and mood. First, we predicted that cognitive abilities will positively correlate with users' view and experience of CAV HMI usability as they might shape older user's attitudes towards automation (Schaefer, Chen, Szalma, & Hancock, 2016). Second, that personality factors such as: neuroticism anxiety, aggressive hostility will correlate negatively with self-reported CAV HMI usability, while activity, sociability and impulsive sensation seeking will correlate positively with their perception of CAV HMI usability. Third, positive attitudes towards computers and trust in technology traits will correlate positively with reported CAV HMI usability. Fourth, presence, trust in automation state will also correlate positively with their perceived CAV HMI usability, while simulator sickness will have a negative correlation. Positive emotions will correlate positively with perceived CAV HMI usability and negative emotions will negatively correlate with their experience.

2. Literature Review

2.1. Designing Human Machine Interfaces for Automated and/or Connected Automated Vehicles

In fully AVs and/or CAVs the human role is likely to change from that of a driver to a passenger as s/he will no longer engage in traditional driving-related tasks (e.g., steering, accelerating, braking, lane changing) with the opportunity to take both eyes and mind off the road (Beggiato et al., 2015; Li et al., 2019). Consequently, the user might choose to engage in other activities (e.g. read a book, watch television, eat in, browse the internet) (Ferati et al., 2017; Li et al., 2019; Morgan et al., 2018; Shergold, 2018; Voinescu, Morgan, Caleb-Solly, & Alford, 2018), but they would still require an HMI to, for example, inform and update them on what is happening (e.g. to query system status and check energy/fuel levels and to obtain information to support situation awareness such as query location on a map) especially when they have been disengaged, (Li et al., 2019). Despite this possible changing role of the human within road vehicles as a controller of many of the systems, it is highly likely that CAVs will require specially designed HMI dashboards to respond to the new needs of users and system challenges. Currently, Level 5 CAVs are quite still far off before deploying them on public roads (Kyriakidis et al., 2019), as the existing technology is still lacking and more technological advances have to be made (Nikitas, Njoya, & Dani, 2019). To cover this gap, researchers have started to explore the design of HMIs for AVs/CAVs, though very few have focused on fully automated driving, and of these still fewer included participants that have actually experienced a highly or fully automated journey(s) (Nordhoff et al., 2016). A summary of existing HMI principles for Levels 3 to 5 automation and type of methodology employed is synthetized in Table 1. Simulator-based studies are needed to enhance current knowledge on Level 5 CAVs, on terms of design and testing of the technology, issues that are crucial to the adoption of AVs/CAVs in the near future (Nikitas et al., 2019).

2.2. Designing Automated and/or Connected Automated Vehicles' Human Machine Interfaces for Older Adults

2.2.1. Older Adults' Cognitive Functions and Their Implications for Automated and/or Connected Automated Vehicles' Human Machine Interface Design

It is well accepted that ageing is associated with a series of changes including psychological (e.g. cognitive domains such as memory, attention), physical mobility as well as sensory (e.g., vision and hearing) decline (Deary et al., 2009; Freedman, Martin, & Schoeni, 2002; Glisky, 2007). Due to variations in ageing-related impairments across individuals, it is recommended that vehicle HMI designers ensure that interfaces are usable by individuals with a diverse range of needs and abilities, including older adults (Fisk et al., 2009; Holzinger, Searle, & Nischelwitzer, 2007; Li et al., 2019; Morgan et al., 2017; Naujoks, Wiedemann, Schömig, Hergeth, & Keinath, 2019). With this in mind, when older adults interact with a system not designed for them (e.g., designed for younger-middle aged people), errors tend to occur more often as reduced (e.g., short-term memory capacity) and/or changing (e.g., ability to switch and/or hold attention) capabilities may not have been considered in the design process. By understanding a range of factors including sensory, cognitive and physical differences of older adults, systems can be better designed to match the usability requirements for the population of interest, including training of cognitive abilities and augmenting or substitution of underlying limitations that can help older adults to benefit most from emerging technologies (Charness & Boot, 2009; Charness, Yoon, Souders, Stothart, & Yehnert, 2018; Holzinger et al., 2007). These factors might impact the acceptance of automation (Schaefer et al., 2016), as cognitive abilities influence the understanding of how automation works and they impact the level of self-perceived ability to use automation, and shape the expectations related to automation (Schaefer et al., 2016).

2.2.2. Current Guidelines for Automated and/or Connected Automated Vehicles' Human Machine Interfaces for Older Adults

Even with research acknowledging the importance of cognitive and sensory difficulties associated with ageing in HMI design, CAV and/or AV HMI design principles for older adults remain under-studied, especially in relation to fully automated Level 5 technology. For example, Morgan et al. (2017) reviewed and synthetized fully automated CAV HMI design principles for older adults without testing participants in a fully AV simulator. In a recent study involving a sample of older adults experiencing driving sessions in highly AV simulator that required hand back control of the vehicle after a session of fully AV driving mode, Li et al. (2019) investigated older participants' attitudes and perceptions towards AV HMIs and provided recommendation for AV HMI design. A summary of findings and guidelines is presented in Table 1.

2.3. The Role of Individual Differences in Automated Vehicles' Human Machine Interface Design

2.3.1. Older Adults' General Characteristics and Automated and/or Connected Automated Vehicles' Human Machine Interface Design

Recent studies, including one using Eurobarometer data from 2014 shows that older adults are less favorable towards AVs than younger individuals (Hudson et al., 2019; Hulse et al., 2018). Men are more willing to travel in an AV/CAV and less likely to worry about automation failure then women (Nordhoff et al., 2016). Women tend to report higher concern with AVs and report less benefits of AVs (Acheampong & Cugurullo, 2019; Charness et al., 2018; Hulse et al., 2018; Schoettle & Sivak, 2014). Age is also a negative predictor of perceived benefits of AVs (Acheampong & Cugurullo, 2019) and middle-aged respondents compared to younger respondents report increased concern with AVs (Charness et al., 2018). Younger adults are more likely to respond that they would ride in AVs and are more interested in adopting AVs (Schoettle & Sivak, 2014). Driver status is not a significant predictor of general attitudes towards AVs (Hulse et al., 2018). Experience and familiarity also seem to impact acceptance of AVs/CAVs as the more someone is familiar with the technology, the more likely they are to accept it (Nordhoff et al., 2016). Prior knowledge of AVs is associated with less concern and being willing to relinquish driving control (Charness et al., 2018) and with positive views towards AVs including crash reduction, fuel economy and less concern about learning to use them (Schoettle & Sivak, 2014). Participants with least knowledge of AVs yield most negative views towards them (Sanbonmatsu, Strayer, Yu, Biondi, & Cooper, 2018) and are more likely to respond that they would not ride in them (Schoettle & Sivak, 2014). Education emerges as a significant predictor of perceived ease of use and benefits of AVs (Acheampong & Cugurullo, 2019) and higher educational status is associated with increased likelihood of having self-driving technology and willingness to ride in AVs (Schoettle & Sivak, 2014). Early adopters of AVs appear to be consumers with high income, good knowledge of AVs and with positive perceptions towards technology in general and AVs, in particular (Hardman, Berliner, & Tal, 2019). Taken together these findings highlight the importance of designing HMIs with user input, and suggest that personal characteristics such as age, attitudes towards technology and direct experience, might have an important role in usability ratings.

Research also reveals that higher levels of new technology use, perceived usefulness and ease of use seem the best predictors of acceptance of AVs, as well as trust in automation (Choi & Ji, 2015; Ekman, Johansson, & Sochor, 2016; Souders & Charness, 2016). Current use of automation-vehicle technology is associated with increased interest in having fully automated technology (Schoettle & Sivak, 2014). Perceived benefits and trust also positively predict willingness to pay for AVs. Perceived risk and dread negatively predict willingness to pay for AVs

(Liu, Guo, Ren, Wang, & Xu, 2019). Similarly, low trust in technology is associated with more negative views and attitudes towards AVs (Sanbonmatsu et al., 2018; Zhang et al., 2019). Favourable attitudes towards technology emerge as significant predictors of perceived benefits of AVs while perceived ease of use predicts affective attitudes towards AVs (Acheampong & Cugurullo, 2019).

Usability might be a key determinant to willingness to use CAVs by the older population. In a broad sense, usability can be defined as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241-11:2018). Recently, Morgan et al. (2017) defined CAV HMI usability as: "…aspects of the HMI including: learnability; efficiency; memorability: error handling; and satisfaction (linked with likelihood of continued use)" (p. 328). Previous research highlights that usability impacts drivers' trust of AVs (Merat, Madigan, & Nordhoff, 2017) and their acceptance of new technologies (Horberry, Regan, & Stevens, 2017; Martens & Jenssen, 2012).

2.3.2. Personality and Automated and/or Connected Automated Vehicles' Human Machine Interface design

The focus of the current paper is on the relationship between individual difference factors and perceived CAV HMI usability. A recent review reported that half of the studies that focused on AVs investigated participants' behavioral characteristics and perceptions that can impact their willingness to use AVs and on desirability of AVs (Gkartzonikas & Gkritza, 2019), noting that most studies were online surveys and not experiments using e.g., simulated or real-life automated journeys. Despite individual differences playing an important role in the very popular Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003), they were not studied thoroughly, and crucially, additional research is needed (Devaraj, Easley, & Crant, 2008; Schaefer et al., 2016). The results of our literature review concerning personality factors and how they impact attitudes towards AV/CAV and HMI design are summarized in Table 2 and highlight that not only has the impact of personality on in-vehicle CAV HMI design and/or interaction received limited attention, but the studies that report significant correlations between personality factors and AV/CAV acceptance, adoption and willingness to use, have not actually tested participants in an AV driving (simulator or road vehicle) situation, but instead were online survey studies (see Table 2). The current study addresses this gap by being the first to expose participants in a simulated CAV environment.

2.4.Situational Factors and Automated Vehicles' Human Machine Interface Evaluation

Usability studies benefit from an ecologically valid evaluation setting. In our case, a virtual reality (VR)-based methodology (Rebelo, Noriega, Duarte, & Soares, 2012) which allows us to investigate the usability of the CAV HMI in a setting that resembles real life situations. There are several variables important for studies that focus on the use of VR and simulators in human behavior because they relate to performance obtained in the virtual world. The first variable is the sense of presence, a feeling of actually "being there" in the virtual environment (Kennedy, Lane, Berbaum, & Lilienthal, 1993). Presence seems to positively influence outcomes studied in VR as it enhances performance (Ai-Lim Lee, Wong, & Fung, 2010; Dinh, Walker, Hodges, Song, & Kobayashi, 1999; Lin, Duh, Parker, Abi-Rached, & Furness, 2002; Price, Mehta, Tone, & Anderson, 2011) and correlates with usability in a virtual environment (Brade et al., 2017). The more participants experience a sense of presence in the virtual environment, the more positively they will rate the usability of the system (Brade et al., 2017). Presence is often described as an essential component of usability and user experience in virtual worlds (Tcha-Tokey, Loup-

Escande, Christmann, & Richir, 2016) which makes it an important variable in simulator and usability studies probably because an increased sense of presence in the virtual environment will make the experience as real as possible which allows for an accurate and ecological usability evaluation (North & North, 2016).

In VR literature, potential negative aspects of such realistic simulations, including simulator sickness are discussed. Simulator sickness describes symptoms similar to motion sickness. Symptoms occur during exposure in virtual environments and include general discomfort, fatigue, headache, eye strain, stomach awareness, nausea, dizziness, vertigo, and burping, sweating, blurred vision (Kennedy et al., 1993). Previous research has shown that simulator sickness symptoms are negatively associated with presence and might negatively impact performance (Kennedy et al., 1993; Maraj, Badillo-Urquiola, Martinez, Stevens, & Maxwell, 2017; Milleville-Pennel & Charron, 2015), for example, interaction with a CAV in-vehicle simulator HMI. Increased symptoms of simulator sickness have been proposed to hinder the adoption of AV on a large scale, as various acceleration and deceleration tasks during simulated journeys are responsible for escalation of these symptoms (Jones et al., 2019). In the current study we also explore whether negative simulator sickness symptoms correlate with self-reported usability, a link that, to the best of our knowledge, has not been tested before in a full VR-based CAV simulator with older adults.

Usability (Lallemand, Gronier, & Koenig, 2015) and CAV acceptance can also be influenced by the mood and emotions of the participants as emotions are considered part of decision-making process (Schwarz, 2000). Based on UTAUT (Venkatesh et al., 2003), Nordhoff et al. (2016) proposed several factors to help understand and predict the user acceptance of pod-like AVs/CAVs, including individual characteristics (e.g. personality, demographics), and participants' mood and emotions, which can impact user experience and usability and acceptance of CAVs. Some consider that user experience and usability is a process which is shaped by the users' response to the use of technology (Jokinen, 2015) and that both positive and negative emotions can influence the willingness to ride in an AVs (Anania, Mehta, Marte, Rice, & Winter, 2018). Therefore, investigating the emotional response during the CAV HMI evaluation would definitely improve our understanding of factors that enhance the perceived usability of a CAV HMI designed for older adults.

3. Method

3.1. Participants

Twenty-five individuals aged 65-83 years old (M = 70.20, SD = 4.46) participated[5]. This provided an adequate sample to detect medium to large effect sizes (Pearson's r = .40 to.45)[6] with power of .80 (Cohen, 1988) on our main measures. A priori power calculations were conducted using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009). A post hoc power analysis using G*Power revealed that for our sample of 25 participants in order to detect large effects of .50 we had a power of .87; for medium to large effects of .40 we had a power of .68; for medium effects of .30 we had a power of .45; and for small effects of .10 we had a power of .12.[7] The sample was male dominant (N = 18, 72.0%).

Inclusion criteria were: fluency in English language and comprehension, age equal to or above 65 years. Twenty-four participants (96%) had corrected vision and four had corrected hearing

[5] For world developed countries the age of 65 years old is a cut-off for defining old age (e.g., Michel, Beattie, Martin, & Walston, 2018; WHO, 2002)
[6] Pearson's r was used as a measure of effect size using the well-known benchmarking criteria: .10, .30, and .50 indicating small, medium and large effects respectively (Cohen, 1988).
[7] Similar sample sizes were used in other AV simulator studies (e.g. Fredrick Ekman, Johansson, Bligård, Karlsson, & Strömberg, 2019; Ritchie et al., 2019; Strauch et al., 2019; Swan, Shahin, Albert, Herrmann, & Bowers, 2019)

(16%). Three participants were also taking antidepressant medication. All participants were highly functional and able to complete the entire study. Exclusion criteria were: the presence of any severe health conditions (i.e. epilepsy, severe neurological impairments, heart surgery, recent major road traffic accident experience), mild to moderate visual and hearing impairments. Participation was voluntary, and each participant received a £20 voucher to cover transport and associated costs. See Table 3 for full participants' characteristics.

3.2.Materials

3.2.1. Overview of experimental set-up.

The perceived usability of the CAV HMI was tested by exposing participants to a fully CAV (Level 5) simulator. The CAV simulator consisted of a) a simulator pod shell where participants were seated, b) a large standard flat-screen computer monitor that displayed the virtual journeys, and, c) four virtual journeys (driving scenarios) that consisted of real-world video footage recorded using GoPro cameras. For the CAV HMI usability evaluation, the participants interacted with the HMI during the four virtual journeys and completed post-experimental questionnaires at the end of the virtual journeys. The HMI was positioned on the CAV simulator dashboard. Both CAV simulator and HMI were designed to mimic a fully CAV: a) the virtual journeys were pre-recorded to display the pre-planned journey of a fully CAV driving mode (e.g. take the passenger from A to B without any input from the passenger such as change lane, reduce speed), b) the HMI displayed the pre-recorded route of the pre-planned journey. The result was that the participants experienced four virtual journeys in an experimental set-up that resembled fully automated driving conditions.

3.2.2. HMI.

The design of the HMI was informed by an extensive literature review (Morgan, Caleb-Solly, Voinescu, & Williams, 2016; Morgan et al., 2017) and public engagement workshops with older adult participants that synthesized best practices and recommendations for the design of HMI for older people (Shergold, 2018). The HMI was designed and developed to be a standalone dashboard of the CAV simulator and was pre-programmed to synchronize with the CAV simulator journeys. It consisted of basic functions potentially useful for CAV journeys (e.g., date, time, destination, vehicle status, navigation map) and relied exclusively on visual modality with touch input but without auditory modality. It was implemented on a 12.9-inch iPad Pro with ED backlit display with iPS technology; retina display; 2732x2048 resolution at 264 pixels per inch, and fingerprint resistant coating (Figure 1-4). A summary of the HMI features can be found in Appendix A Table 1.

3.2.3. Simulator Setup.

Simulator shell.

The simulated journeys were experienced inside a static Lutz pod shell (Figure 5 and Figure 6). The Lutz pod was designed and supplied by Transport System Catapult (TSC). It is a two-seater pod with 2 doors. The steering wheel was removed, as we were simulating fully automated journeys with no human input to driving components following journey set up. Participants could engage with interactive features of the HMI (Figure 1-4) but could not modify things like journey destination, route-to-destination, or driving behaviour (e.g., speed, style) once the initial journey had been set up[8].

[8] The current version of CAV HMI is the first of a number of studies forming part of a large multi-partner project and was developed using an iterative design process.

Simulator Integration/Virtual Environment.

We used a non-immersive VR set-up (Rizzo & Koenig, 2017) as the content was video-based and delivered on a large standard flat-screen computer monitor. The CAV VR simulator consisted of a static, fixed position Lutz pod shell and a large screen Samsung 60 inch 4K Ultra HD Smart TV 3840x2160 pixels sited in front of the pod where the virtual journeys were displayed (Figure 7).

The driving scenarios were recorded using GoPro Hero 4 cameras. The cameras were mounted on a 3D printed static support specially designed and built by XXXX XXXX XXX personnel to be robust, easy to attach without damaging the bonnet of the car and that allowed recording at speeds of up to 40 mph. The recordings were performed while driving a Hybrid Mitsubishi Outlander that largely mimics the view that would be perceived from the Lutz pod. The driver had over 20 years of driving experience and no penalty points. The driving style was cautious and defensive (e.g., anticipated potential hazards, avoided risky manoeuvres such as driving through amber lights, drove at a safe speed and distance from other vehicles) with another two researchers in the car with the navigation system and looking for hazards. To emulate an electric CAV and increase the similarity of real-world situations of future CAVs, the Outlander was driven mainly in electric mode during recording. GoPro cameras mounted externally on the front of the bonnet did not pick up the internal engine noise (either electric or petrol) but did record the external surroundings to provide an authentic journey experience. Noting that recent literature points out that future AVs and CAVs will be mostly electrical for environmental reasons, which is also likely to increase the public acceptance and adoption of CAVs (Webb, Wilson, & Kularatne, 2019; Wu, Liao, Wang, & Chen, 2019; Yi, Smart, & Shirk, 2018).

In total, four journeys were selected, filmed mainly on public roads in and around XXXX XXXX XXX, during the summer months and gave a widespread representation of driving within inner and outer urban XXXX XXXX XXX settings, with a mixture of road types with speed limits ranging from 20-40-mph, a mixture of road infrastructure (e.g., traffic lights, crossings, roundabouts), and varied backdrops (from highly built-up to outer city suburbs with green spaces such as parks). Each route was carefully planned using Google Maps and an AA Route Planner (http://www.theaa.com) and included a 2-minute stop halfway though. The total time of the journeys was approximately (±30-seconds) 7 minutes, with a 2-minute stop, and a 5-minute drive.

The four virtual journeys experienced by the participants were:

a) Railway Station (starting point) to Medical Centre (intermediate stop) to Home (destination);

b) Home (starting point) to Dental Clinic (intermediate stop) to Hospital (destination);

c) School (starting point) to Park (intermediate stop) to Leisure Centre (destination);

d) Gym (starting point) to Public House (intermediate stop) to Home (destination).

During the virtual journeys, participants were seated in the Lutz pod (recommended to sit in the centre of the bench seat – though wide enough for two people) and the virtual journeys were displayed on the 60-inch Samsung TV screen via an Alienware I7, 2.60GHz processor laptop with a resolution of the display at 1920×1080 at a refresh rate of 60 Hz. See Figure 7 for an example the simulator virtual environment/simulator/HMI integration.

3.3.Measures

Several psychological scales, tests and questionnaires were administered pre-journey, during journey and post-journey or both (full description of the scales and measures used can be found in Appendix A Table 2).

Self-screening questionnaire. To apply our eligibility criteria, participants self-declared any major health-related conditions (e.g. stroke, epilepsy, heart conditions, recent major road traffic accident experience). To avoid any risky situations where exposure in the CAV simulator might result in a health risk or unpleasant experience to the participants, participants with prior major health-related conditions such as those described above were excluded.

Demographic questionnaire. To collect data on: age, gender, qualifications, marital and occupational status, and current medication.

Cognitive functioning. For cognitive functioning we used several measures: Ospan (Turner & Engle, 1989; Unsworth, Heitz, Schrock, & Engle, 2005), Trail Making A & B tests (Reitan, 1958) and Corsi Blocks test (Corsi, 1972)

Personality. Zuckerman-Kuhlman Personality Questionnaire (ZKPQ-50-CC, shortened form, Aluja, García, & García, 2002; Aluja et al., 2006) was used to measure five personality traits: impulsive sensation seeking, aggression hostility, activity, and neuroticism anxiety.

Trust in technology trait. We used the General Trust in Technology Scale (GTS, Mcknight, Carter, Thatcher, & Clay, 2011).

Attitudes towards computers. Attitudes Toward Computers Questionnaire (ATCQ, Jay & Willis, 1992) was used as a measure of attitudes towards computers.

Simulator sickness. To measure simulator sickness, we used the Simulator Sickness Questionnaire (SSQ, Kennedy et al., 1993).

Presence. We used the Presence Questionnaire (PQ, Witmer & Singer, 1998) to measure perceived sense of presence in virtual environments.

Mood. The Positive and Negative Affect Schedule was used to measure emotion (PANAS, Watson, Clark, & Tellegen, 1988).

Trust in automation state. The checklist for Trust between People and Automation (CTBPA, Jian, Bisantz, & Drury, 2000) was used to measure system dependability and reliability.

System usability. The system Usability Scale (SUS, Brooke, 1986) was used to gain feedback on perceived usability.

3.4.Procedure

The Study was approved by the University XXXX XXXX XXX Ethics Committee which also included a linked and approved risk assessment. Written consent from the participants was obtained prior to the experiment and after they had received and understood an information sheet.

The study had four phases: an In-Vehicle Participant Workshop, followed by a pre-test, test, post-test, and follow-up (with several scales completed at home after the study).

3.4.1. In-Vehicle Participant Workshop.

Each participant had attended an introductory workshop that aimed to inform participants about the goals and objectives (not specific studies, manipulations or predictions) of the XXXX XXXX XXX research project, timescale and to gather information relating to e.g., expectancies about CAVs (e.g., design, journey types, likelihood of using, HMI design features) and clarify the terminology (e.g. AV and CAV and Level 5 of automation). The workshops also served as part of an iterative process to design the HMI (different versions throughout the project), as well as to ensure that all participants' expectancies about automated vehicles were at a similar level to avoid bias through different personal understanding of the project scope and media coverage of the topic of CAVs.

The workshops included an approximately 25-minute session where each participant was able to contribute individually, and as a group. The data being collected was primarily qualitative and concerned three major themes: a) general attitudes towards CAVs (e.g. 'Some people think that cars that will be able to drive themselves will be on our roads in the near future. What do you think about that?'); b) characteristics and functionality of CAVs (e.g., 'What features and characteristics would you want a driverless vehicle to have if you were going to use one, and would that be any different to vehicles now?'); c) type of journeys in a CAV (e.g., 'If you personally had access to a driverless vehicle, what type of journeys might you make and where would you go in it?'). Key findings were: the HMI should be easy to use, clear and robust; preferably all HMIs should have a standard design/approach, so it would be easy to use with various CAVs, avoid using jargon, or computer-speak. Participants preferred adaptability features for a degree of flexibility in the interface to suit different users and the use of large icons, fonts, with labels and pictures, and favour text to icons. Controls should be easy to reach, the screens and controls large enough to be seen without glasses. Physical button for ON and STOP and ability to stop car, get help or go to safe place were also listed as important features. In terms of functions, older adults also suggested the capacity of the HMI to show routes during journey, alternative routes as well, local routes and short cuts, road blockages and any route updates, including hazard conditions. Maps were considered important, to be able to see where the CAV is heading and points of interest.

3.4.2. Pre-test.

After participants arrived at the University facility designated for the automated driving experiments, the experimenter provided them with an overview of the testing session. The first scales administered were: the self-screening questionnaire to assess eligibility (noting administered first followed by informed consent), a demographic questionnaire, and a pretest version of the SSQ and PANAS. All measures were administered in paper-and-pencil format.

3.4.3. Test.

This also known as journey phase. Participants were seated in the simulator and experienced virtual journeys while interacting with the HMI. Four journeys were used. The first was a practice in which participants had a chance to familiarize themselves with the simulator and journeys. The other three were used for the usability assessments and the journey order was counterbalanced using a Latin square with six iterations: ABC, ACB, BAC, BCA, CAB, CBA (equally balanced across the 25 participants). A printed overview instruction leaflet that provided a summary of the test phase was given to participants by the researcher prior to beginning the journeys. Written instructions for each journey were also offered by the researcher to participants in a printed format prior to each journey. The participants were encouraged to read carefully the instruction sheet in order to understand and familiarize themselves with the procedure and tasks. Thus, they kept the instruction sheet during each journey in case they needed to access the information provided. This was chosen to reduce potential bias associated with poor memory retention and reduce cognitive load. Details of the procedure for each journey are presented below.

At the beginning of the first journey, the researcher helped the participant to be seated in the simulator. Then the researcher gave the participant the journey familiarisation instruction sheet that contained an overview of the tasks to be performed during the journey, including familiarisation with the Lutz pod and simulator, instructions for the first journey and what to do in case they got motion sickness symptoms. Any queries were answered by the researcher to ensure that the participants understood the tasks. During the first journey, the researcher sat next to the participant in the pod and set up the HMI, while the participants were free to look, but not touch the HMI. After the familiarization journey, the participant received an instruction sheet before each of the three journeys that explained the tasks for that journey.

For the next three journeys, participants were seated alone in the simulator and read the journey instructions on the sheets provided by the researcher. The journey instruction sheet contained information about:

a) Journey scenario: e.g., “Imagine that you have just picked up your grandchild from school. You will then go to XXXX Park Avenue with your grandchild to buy some ice-cream. After that you will go to XXXXX leisure center together)”
b) Journey duration: e.g., “This journey will last for 7 minutes, including a 2-minute stop at XXXX Park Avenue)”
c) Journey set-up instructions: e.g., “Set up the Destination and Stop; Journey Destination: XXXX centre- XXXX; Journey Stop: XXXX”. To set-up the destination the participants had to select from the HMI menu the destination of the current journey (four destinations options were available, that corresponded to four journeys they had to make in the simulator). After selecting the destination, the participants had to press the start button when they were ready to start the journey. After pressing the start button the journey started.
d) Tasks they have to perform during the journey: e.g., “During the journey, we would like you to check if the vehicle is running satisfactorily. So, try and do the following: 1. Check the status of the vehicle. 2. Check the battery”. These tasks enable a controlled assessment of interaction with the HMI and participants were able to complete the tasks during the journey whenever they chose to.
e) After clarifying any issues, and when the participant was ready, they pressed the start button and the journeys began.

3.4.4. Post-test.

After the journeys ended, participants disembarked from the pod and completed the following scales in paper-and-pencil format: the post journey SSQ, CTBPA, SUS, PQ, followed by the computerized cognitive tests. The cognitive tests were administered in randomized order across participants. At the end of the testing session the participants received the voucher and a pre-paid envelope with psychological scales and questionnaires to be completed at home within 48 hours. We chose to administer these scales and questionnaires at home because they do not require special conditions for administration and are not task dependent and helped to limit the total duration spent at the assessment facility and possible fatigue (over 2.5 hours for each participant). The four scales used for this purpose were: GTS, ATCQ and ZKPQ.

The testing session within the laboratory lasted for approximately 2 hours and 45 minutes, with variability depending on inter-subject individual differences, such that some participants took up to 4-hours to complete the study. Noting that breaks were also offered throughout the experiment at the request of the participant, and many took these opportunities, albeit at various points depending on needs and requirements (e.g., use of toilet, refreshments, rest).

3.5. Design and Data Analysis

To investigate possible associations between variables, a cross-sectional design was incorporated, and Pearson *r* parametric correlations were conducted. We also employed Pearson's r correlations as a measure of effect size (Field, 2009), with effect size showing the magnitude and strength of the relationship between the two variables given the sample size (power). A value of .10 indicates a small effect size, .30 a medium effect size, and .50 reflects a large effect size (Cohen, 1988). Data analysis was performed using *IBM SPSS Statistics for Windows, Version 24.0*. (Armonk, NY: IBM Corp).

4. Results

The aims of the current study were to explore the perceived usability of a CAV HMI designed for older adults by quantifying the existence, direction and magnitude of the correlations between their usability ratings and individual, as well as situational factors.

4.1. Perceived HMI Usability and Individual Differences

As hypothesized, there was a moderate to large relationship between cognitive performance on most measures and CAV HMI self-reported usability ratings (all outcomes measured by Corsi Blocks Test & Ospan, and two out of six outcomes measured by Trails A & B). The general trend in results can be described as the more cognitive performance increases, so does the reported usability of the HMI (see Appendix A Table 3 for Pearson *r* correlation coefficients for HMI usability ratings and cognitive abilities and Design and Data Analysis section provides effect size boundaries).

Working memory (verbal and visual/spatial) (Corsi Blocks Test & Ospan) was found to be strongly positively associated with self-reported usability (SUS) with effect sizes ranging from $r = .44$ to $r = .59$, suggesting that better working memory performance is associated with better experienced usability of the HMI. A moderate to strong negative relationship between executive function (Trails B) reflecting greater completion time and poorer usability (SUS) was reported on two ($r = -.45$, $r = -.48$) out of three Trails B outcome measures, which suggests that better executive functioning performance (faster time to complete task, but not number of correct responses) is associated with increased perception of AV HMI as usable. Older participants that reported increased self-reported usability, were more likely to have better visual search speed, scanning, speed of processing, mental flexibility, as measured by Corsi Blocks Tests and Ospan (Appendix

A Table 3). There was no significant relationship between Trails A and self-report usability ratings, probably because Trails A compared to Trails B reflects less complex cognitive processes being less sensitive to detect impairment of executive functions compared with Part B (Crowe, 1998).

Contrary to our prediction, for personality, results point to a moderate negative relationship only between neuroticism anxiety (ZKPQ) and perceived CAV HMI usability (SUS) (r = -.37) (Appendix A Table 4), while other personality factors were not associated with usability. Participants who are less anxious and neurotic are more likely to evaluate the HMI as better in terms of usability. Contrary to our expectation, other personality factors, like sociability, activity, aggression hostility and impulsive sensation (ZKPQ) seeking did not correlate with the quality of usability of the CAV HMI (Appendix A Table 4 displays Pearson r correlation coefficients for HMI usability ratings and individual differences and Design and Data Analysis section provides effect size boundaries).

Finally, and as predicted general trust in technology (GTS), as a stable trait, was also positively and moderately associated with high usability scores for the HMI (SUS) (r = .47), although state trust in automation (CTBPA) was not (Appendix A Table 4 & 5). Contrary to our expectation, attitudes towards computers and state trust in automation are variables that did not relate to usability of the CAV HMI (Appendix A Table 4).

4.2. Relationships Between Usability of the HMI and VR Simulator Context Variables

In line with our prediction, there was a moderate negative association between users' rating of usability of the CAV HMI (SUS) and simulator sickness (SSQ) (r = -.37). We also report a non-significant association between perceived usability of the CAV HMI (SUS) and presence in the simulator (PQ) (Appendix A Table 5). Pre-test post-test simulator sickness (SSQ) comparisons revealed increased post-test symptoms within accepted range (detailed data analysis and results are presented in Appendix A). Contrary to our expectations, trust in automation (CTBPA), in our study directly referred to as trust state in the current CAV HMI was not significantly related to their perceived usability (see Appendix A Table 5 for Pearson *r* correlation coefficients for HMI usability and presence, simulator sickness, and trust in automation and Design and Data Analysis section provides effect size boundaries).

Table 6 (Appendix A) displays Pearson *r* correlation coefficients for CAV HMI usability and positive and negative emotions measured by PANAS. Thus, as expected, there was a moderate to strong positive relationship between positive emotions (PANAS) and perceived CAV HMI usability (SUS) for all four journeys (*r* coefficients ranging from .38 to .59). Results also indicated strong negative correlations between negative emotions (PANAS) and same usability for two out of four journeys (r = -.53, r = -.58), which indicates that for Journey 3 and 4 (last two journeys) reduced negative emotions associated with increased HMI self-report usability. There were no differences across journeys between positive and negative emotions and self-report HMI usability (a detailed data analysis can be found in the Appendix A).

5. Discussion

We investigated the relationship between perceived usability of a fully CAV HMI designed for use by an older adult population, and individual differences, including: cognitive abilities, personality, attitudes towards computers and trust in technology and situational variables like: simulator sickness, presence and mood. The findings highlight the correlation between working memory (verbal and visual/spatial) and perceived usability, for a sample of older adults. The higher the level of cognitive ability, the higher the perceived CAV HMI usability. Similar to the recommendations from the literature concerning HMI design for older adults, which stress the importance of an HMI design to help overcome cognitive difficulties associated with ageing, our study shows that increased perceived HMI usability was associated with higher levels of cognitive ability (Charness & Boot, 2009; Fisk et al., 2009; Morgan et al., 2018; Souders & Charness, 2016). Also, as Schaefer et al. (2016) demonstrated, it might be the case that high cognitive abilities improve participants' understanding of automation and how to use it, which in our study was reflected in better usability ratings for those with increased working memory (verbal and visual/spatial). Following the principles synthesized by Morgan et al. (2017, 2018) and using feedback from older participants during initial workshops on CAVs for participants, our CAV HMI was designed to ensure that it was simple, easy to use, with minimal clutter, with items organised in a natural and consistent way, and had zoom-in and touchscreen capabilities. Overall, our results might inform future AV HMI designs for older adults (see Morgan et al., 2017 for a full review). There is encouraging evidence in favour of simple and easy to use HMIs that do not require a great amount of cognitive resources, especially working memory.

To our knowledge, the study presented here was the first that looked into possible ways in which personality and individual differences relate to the perceived level of CAV HMI usability, and most importantly, it did this by assessing older participants experiencing a CAV simulator, thus increasing the ecological validity of our results compared to survey based studies. This is crucial, as individual differences are important in the UTAUT (Venkatesh et al., 2003). One study conducted by Devaraj et al. (2008) suggests that it is highly likely that a user who scores high on conscientiousness and agreeability, is emotionally stable and extrovert, is more likely to use technologies. Similarly, extraversion, consciousness and emotional stability are associated with behavioral intention to use, and openness to experience and positively related to perceived ease of use (Svendsen, Johnsen, Almås-Sørensen, & Vittersø, 2013), while a risk taking trait also is associated with technology acceptance (Dillon, 2001; Wang, Vang, Lookadoo, Tchernev, & Cooper, 2014). Conscientiousness positively predicts concern with AV, and negatively predicts eagerness to adopt AVs. Emotional stability and openness positively predict eagerness to adopt AVs as well. Extraversion and openness positively predict willingness to relinquish driving control (Charness et al., 2018).

In our study we identified a negative correlation between neuroticism anxiety and usability, which indeed, appears to be mostly related to a good HMI self-reported usability, as described in the above studies. In our study, sociability seems not to correlate with perceived CAV HMI usability. However, in the Alternative Five-Factor Model, which is measured by the ZKPQ questionnaire, compared to the Five Factor Model (Costa Jr & McCrae, 1992; Costa & McCrae, 1992) the impulsive sensation seeking factor also contains some extraversion trait components as described in the factor structure (Sârbescu & Neguţ, 2013). This means that to some extent, some facets of extraversion also related to better AV HMI usability experience as found in our study. Contrary to Nordhoff et al. (2016) who suggest that sensation seekers are less likely to accept automation because handing over control to the AV might reduce the thrill and excitement of driving, we found a non-significant correlation between sensation seeking and AV HMI usability ratings. The hypothesized link between the novelty of experiencing CAV journeys did not obtain support, noting that post hoc power analysis for this correlation revealed that our study was underpowered (achieved power of .49) to detect moderate effect sizes (r = .32). The trend in our results, even if not significant, is similar to that reported by Payre, Cestac, and Delhomme (2014) for AVs acceptance, who point out that sensation seekers evaluate HMI usability more positively because, in the first instance the AV and HMI experience is novel and thrilling, and if they don't get the chance to get used to it, will not become bored. In our recent research[9] we investigated the relationship between personality and preferences of older adults for functions embedded in an in-vehicle CAV HMI. Our results suggest some influence of personality traits measured, as well the Alternative Factor Model (Aluja et al., 2006), with less anxious and neurotic participants displaying preferences for more functions, and participants high on aggressive hostility having reduced preferences for AV HMI functions (Voinescu et al., 2018). However, it should be noted that in this previous study, the CAV HMI interaction and evaluation took place without the exposure of participants to CAV simulated journeys. Again, it might be the case that during a simulated CAV journey, participants might rate their preferences differently to how they would normally do, when the evaluation takes place without interaction during the simulated journeys.

In opposition to findings that highlight the relationship between trust in automation and user experience of AVs (Ekman et al., 2016; Mirnig, Wintersberger, Sutter, & Ziegler, 2016), our study also shows out that state trust in automation in the CAV simulator did not correlate with HMI usability scores, but general trust in technology did. In a recent Level 4 AV simulator-based study with younger and older adults, Molnar et al. (2018) identified that prior use of novel technologies did not predict trust in AV technology, but trust in AV technology predicted engagement with this technology. Our results are in line with those from a number of recent studies that also point out that general trust in technology is one of the key variables in acceptance and positive attitudes towards AVs (Acheampong & Cugurullo, 2019; Liu et al., 2019; Sanbonmatsu et al., 2018; Zhang et al., 2019) which also predicts intended Level 3 AV use (Buckley, Kaye, & Pradhan, 2018a). This means that the older participants in our study who show an increased trait

[9] Voinescu et al. (2018) tested part of the same sample of participants as the current study. Noting that within the Voinescu et al. study a different HMI was evaluated, post simulator journeys, and the study had different objectives and research questions.

of trust in technology, but not state trust in the current CAV HMI, are more likely to evaluate the HMI as more usable. One possible explanation might be the fact that even if our older participants display trust in technology, they might have difficulties trusting the VR CAV simulator, as previous studies have shown that user experience seems to be influenced by trust in automated systems (Ekman et al., 2016; Mirnig et al., 2016). Similarly, Manawadu, Kamezaki, Ishikawa, Kawano, and Sugano (2017) found multimodal HMIs in a level 4 AV reduced self-perceived driver workload, improved the efficiency of interaction measured by input time, and reduced input errors compared with unimodal HMIs. However, in the above studies, the sample consisted of healthy younger adults compared to our study, which tested older adults. Our CAV HMI did not have speech capabilities as previous public engagement workshops with older adult participants revealed that they prefer simple and easy to use CAV HMIs (see Morgan et al., 2016 for a full review) and because the HMI was at its first iteration.

As previous studies suggest, lack of familiarity and lack of previous experience and/or knowledge of AVs may be related to trust in AVs and positive attitudes towards AVs (Acheampong & Cugurullo, 2019; Charness et al., 2018; Hardman et al., 2019; Sanbonmatsu et al., 2018). For example, the novelty of the situation and the fact that our VR CAV simulator was a pod-like vehicle without a steering wheel and pedals, might have contributed to the decrease in trait trust (Nordhoff et al., 2016). Contrary to our expectations, attitudes towards computers was not associated with self-reported CAV HMI usability. It might be the case that our sample of older participants consisted of highly educated people with previous computer experience, which in turn might have improved their level of attitudes and user behavior (González, Ramírez, & Viadel, 2015). Almost 50% of our sample held a University degree, and education appears to be a predictor of perceived ease of use and benefits of AVs (Acheampong & Cugurullo, 2019). It is important to note that we did not measure previous computer experience directly, but all of our participants had previous experience in interacting with touch screen devices, which might account for good levels of computer self-efficacy.

When it comes to simulator sickness, we identified that a reduced level of simulator sickness correlates with usability. As our older participants reported low levels of simulator sickness, the more likely they were to report increased levels of AV HMI usability. Indeed, previous research has also shown that simulator sickness is negatively associated with presence and impacts negatively task performance (Kennedy et al., 1993; Maraj et al., 2017; Milleville-Pennel & Charron, 2015), thus, older participants (aged over 60 years old) in a simulated journey reported lower levels of simulator sickness than younger adults (Jones et al., 2019). Such a result clearly suggests that user experience and usability studies in VR simulator researchers should pay attention to simulator sickness and make sure to deliver virtual reality-based simulator experiences free of simulator sickness to avoid its' negative impact on usability ratings and performance. Contrary to previous studies which identified a positive relationship between presence and usability and performance in a virtual environment (Brade et al., 2017; Dinh et al., 1999; Lee, Wong, & Fung, 2010; Lin et al., 2002; Price et al., 2011) our study failed to detect a significant correlation between the two variables, though it reflects a clear tendency to significance ($p = 0.055$) which might suggest the development of further immersive virtual environments that clearly replicate real-world environments and can be well used for usability testing. However, another study that assessed if realism of the simulated AV driving experience (Level 4 AV automation) can predict engagement with AV technology concluded that the association is not significant (Molnar et al., 2018). Despite the fact that authors used a single question to assess the realism of the AV experience and because presence is a more complex construct, it might be the case that the realism and a sense of being part of the environment do not directly impact AV experience, thus future studies with larger samples of participants should investigate this.

Our study also revealed that there are no differences across journeys in terms of positive and negative emotions reported by our older participants, but older participants that experienced increased positive and reduced negative emotions evaluated the AV HMI as more usable. Though the importance of mood and emotions in HMI design and user experience and usability is stressed in a few studies (Anania et al., 2018; Jokinen, 2015; Nordhoff et al., 2016), we believe our study is the first one that investigated the relationship of emotions and HMI among older participants, and more important in a VR CAV usability evaluation. Anania et al. (2018) testing a sample of younger adults point out that both happiness and fear influence the willingness to ride in an AV. Similarly, Buckley, Kaye, & Pradhan (2018b) identified that reported feelings of relaxation and enjoinment among younger individuals during a Level 3 AV journey were correlated with general trust. Together with the results of our study, this makes the measuring of emotional response to the automated simulator an important, but under-studied parameter for including in HMI usability studies. Increasing elements that induce positive emotions during the usability interaction and reduce negative emotions can account for better usability.

Finally, our research is in partial agreement with UTAUT (Venkatesh et al., 2003) which proposes four major predictors of user acceptance: performance expectancy, effort expectancy, social influence, and facilitating conditions. The utility of applying the UTUAT to AVs was synthetized by Nordhoff et al. (2016) and highlights the importance of individual characteristics (e.g. demographics, personality). In our study, we focused on variables that are associated with usability, which might be a key determinant to willingness to use CAVs by the older population (e.g. Morgan et al., 2017), as usability impacts drivers' trust of AVs (Merat et al., 2017) and their acceptance of new technologies (Horberry et al., 2017; Martens & Jenssen, 2012). As proposed in UTUAT, individual factors play an important role in user acceptance. In light of UTUAT which posits that age is a moderator of behavior intention and usage of new technology, we propose that underlying cognitive factors associated with ageing can account for user acceptance (e.g. a decline in cognitive functioning across older individuals may impact their ability to understand and implement new technology), as in our study, most reduced scores on cognitive functioning were associated with low usability. Despite the fact that it was proposed that individual differences such as personality in the acceptance of AVs (see Nordhoff et al., 2016 review on UTUAT and AVs) might impact the acceptance of AVs, in our study, only neuroticism anxiety and general trust in technology were associated negatively, and respectively positively with self-report usability. In the original UTAUT model, anxiety influences usage behavior, which we also observed in our study. In addition to the UTUAT framework applied to AVs proposed by Nordhoff et al. (2016), we propose that for studies that use CAV/AVs simulators, situational factors may also play an important role in positive ratings of self-report usability namely positive emotions and reduced simulator sickness, both can be accountable for reduced anxiety, which is being acknowledged as important in the UTAUT model.

6. Limitations and Future Directions

We need to mindful about the challenges of introduction of AVs/CAVs on the public roads, especially Levels 4 and 5 of automation. If Level 2 and close to Level 3 AVs have been deployed on public roads, Level 4 and 5 AVs are still likely to be way off, in terms of mass deployment and commercialization (Kyriakidis et al., 2019) with some researchers doubting that fully AVs will be available on public roads (Shladaver, 2016). Rigorous research is crucial to ensure that fully AVs operate at acceptable levels and that elements are designed optimally for particular population sectors (e.g., vulnerable road users). In respect to this, our study might help inform proper HMI AVs design for older population.

Our study has several limitations. First, our sample was underpowered for regression and moderation analysis as we only had a set-time to run the study due to project goals and interdependencies between multiple work streams and partners. Most importantly, we included a sample of a vulnerable population, that was tested in a novel VR CAV simulator. Due to very long testing time per participant and ethical consideration related to this, we decided to go on with a cross-sectional correlation design. As a post-hoc analysis of achieved power showed our sample size had reduced statistical power to detect small effects. Most importantly, we included a sample of a vulnerable population, that was tested in a novel VR CAV simulator. Due to very long testing time per participant and ethical consideration related to this, we decided to go on with a cross-sectional correlation design. Despite this, we provided a set of valuable variables that expand our current understanding of impact on user characteristics on perceived AV usability. Future studies might further investigate the role of individual differences and situational factors and draw conclusions based on which of the variables have the most impact on perceived usability, and also, look into possible moderators or mediators. Second, our study had a cross-sectional design which did not allow us to investigate a cause-effect relationship and run comparisons across various CAV HMI interfaces. In the future, it will be important to test whether the perceived usability of CAV HMIs that differ in terms of e.g., features and functions, is influenced by individual and situational factors (e.g. if functionally simple CAV HMIs are preferred over functionally complex HMIs by older adults that face cognitive difficulties, or if adaptive HMIs work better). As our participant engagement workshops revealed a preference of older participants for simple CAV HMIs and because this was the first iteration of the CAV HMI (another three versions planned during the project) we aimed for a basic CAV HMI. This was also to reduce possible confounds of adding too much complexity. Consequently, we did not include multimodal (e.g. voice interaction) and/or adaptive capabilities besides basic iPad settings (e.g. changing layout), despite the literature supporting its inclusion, but in future planned studies we will explore further this hypothesis. In the current study the journey set-up was predefined, with limited interactivity, mostly due to pre-recorded 360° video capabilities which can increase realism, but reduce interactivity compared to computer-generated virtual environments which can be easily manipulated and integrated within the simulator. The journeys were restricted to urban/city surroundings, all in clear visibility, with good weather which we hypothesize would favourably impact usability. It might be the case that other weather conditions (e.g. fog, heavy rain), night-time driving conditions and rural surroundings could impact usability ratings, but these hypotheses are planned by us for future experiments.

To measure perceived HMI usability, we employed a subjective measure and did not include other objective measures (e.g. HMI interaction data, eye-tracking data), but these are planned for future HMI iteration studies. Our sample of participants consisted mostly of older adults with a good level of cognitive functioning and did not include participants diagnosed with mild cognitive impairment, and neither participants with severe visual and hearing difficulties. We wonder how individual differences among these samples of participants might impact experience and design of CAV HMIs.

7. Conclusions

The current study aimed to investigate the relationship between individual differences and situational context and the perceived usability of a fully CAV HMI designed for older adults. Our work is highly relevant for the study of CAV HMIs for older adults, noting the conclusions drawn are based on VR CAV simulator exposure, with a CAV HMI designed for older adults, which makes the assessment context immersive, adding to the novelty and impact of the study. We acknowledge that future research is needed to gather evidence in favor of the use of AVs and CAVs on public roads, and a very good and in-depth understanding of the field. Our study can be viewed as a very important step, which among other current and future studies, can lead to a better understanding of this challenging area of development.

Overall, the findings might suggest that HMIs that are perceived as simple to use and require less interaction, are likely to be preferred by older adults. Perceived HMI usability is closely related to the level of cognitive performance (most components of cognitive processes such as working memory), better cognitive performance is associated with better CAV HMI usability ratings. In terms of individual differences, our results indicate that some individual characteristics such as trust in technology and neuroticism anxiety appear to be related to usability. This means that older adults who trust technology, are less neurotic and less anxious are likely to report a better CAV HMI usability in our trial. Based on our results, we are in favour of using virtual reality as a simulated AV experience to conduct usability evaluations. We mention that careful attention should be paid to reducing simulator sickness and increasing the feeling of presence in a virtual environment.

These findings reinforce the requirement to consider a wide range of individual differences in controlled assessments of CAVs and CAV HMI research. Whilst often not a key research question in itself, personality and other traits can impact on assessment outcomes in otherwise carefully controlled studies. Cognitive factors such as working memory (verbal and visual/spatial) are of specific importance to some relevant populations including the elderly who will be an important sector for future CAV use. These findings point to the clear need for adaptable and flexible interfaces for CAVs if they are to be inclusive to the majority of our aging populations. Importantly, HMIs that are difficult to use may of themselves preclude the take-up of CAVs by this important future user group. These factors therefore need to be borne in mind in planning future studies and assessing outcomes.

Acknowledgement

The research is part of an Innovate UK research project FLOURISH: Empowerment through Trusted Secure Mobility (2016–2019: http://flourishmobility.com/). We thank our collaborators at Designability (Jessica Ridgers, Keir Haines, Hazel Boyd) for designing the HMI and at TSC (Alan Nettleton, Alan Peters, and Roland Guichard) for programming the prototype HMI.

References

Acheampong, R. A., & Cugurullo, F. (2019). Capturing the behavioural determinants behind the adoption of autonomous vehicles: Conceptual frameworks and measurement models to predict public transport, sharing and ownership trends of self-driving cars. *Transportation Research Part F: Traffic Psychology and Behaviour, 62*, 349-375. doi:10.1016/j.trf.2019.01.009

Ai-Lim Lee, E., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education, 55*(4), 1424-1442. doi:https://doi.org/10.1016/j.compedu.2010.06.006

Aluja, A., García, Ó., & García, L. F. (2002). A comparative study of Zuckerman's three structural models for personality through the NEO-PI-R, ZKPQ-III-R, EPQ-RS and Goldberg's 50-

bipolar adjectives. *Personality and Individual Differences, 33*(5), 713-725. doi:10.1016/j.paid.2006.03.001

Aluja, A., Rossier, J., García, L. F., Angleitner, A., Kuhlman, M., & Zuckerman, M. (2006). A cross-cultural shortened form of the ZKPQ (ZKPQ-50-cc) adapted to English, French, German, and Spanish languages. *Personality and Individual Differences, 41*(4), 619-628. doi:10.1016/j.paid.2006.03.001

Anania, E. C., Mehta, R., Marte, D., Rice, S., & Winter, S. R. (2018). *Which Factors Predict Consumer Willingness to Ride in Driverless Vehicles?* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting. doi:10.1177/1541931218621314

BBC News. (2017). Hammond: Driverless cars will be on UK roads by 2021. Retrieved from http://www.bbc.co.uk/news/business-42040856

Beggiato, M., Hartwich, F., Schleinitz, K., Krems, J. F., Othersen, I., & Petermann-Stock, I. (2015). *What would drivers like to know during automated driving? Information needs at different levels of automation.* Paper presented at the 7th Conference on Driver Assistance, Munich.

Bennett, R., Vijaygopal, R., & Kottasz, R. (2019). Willingness of people with mental health disabilities to travel in driverless vehicles. *Journal of Transport & Health*, *12*, 1-12. doi:10.1016/j.jth.2018.11.005

Brade, J., Lorenz, M., Busch, M., Hammer, N., Tscheligi, M., & Klimant, P. (2017). Being there again – Presence in real and virtual environments and its relation to usability and user experience using a mobile navigation task. *International Journal of Human-Computer Studies, 101*, 76-87. doi:https://doi.org/10.1016/j.ijhcs.2017.01.004

Brewer, R., Garcia, R. C., Schwaba, T., Gergle, D., & Piper, A. M. (2016). *Exploring traditional phones as an e-mail interface for older adults.* Paper presented at the ACM Transactions on Accessible Computing (TACCESS), New York, NY.

Brooke, J. (1986). *System usability scale (SUS): a quick-and-dirty method of system evaluation user information*. Reading, UK: Digital Equipment Co Ltd.

Buckley, L., Kaye, S. A., & Pradhan, A. K. (2018a). Psychosocial factors associated with intended use of automated vehicles: A simulated driving study. *Accident Analysis & Prevention, 115*, 202-208. doi:10.1016/j.aap.2018.03.021

Buckley, L., Kaye, S. A., & Pradhan, A. K. (2018b). A qualitative examination of drivers' responses to partially automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, *56*, 167-175. doi:10.1016/j.trf.2018.04.012

Charness, N., & Boot, W. R. (2009). Aging and Information Technology Use:Potential and Barriers. *Current Directions in Psychological Science, 18*(5), 253-258. doi:10.1111/j.1467-8721.2009.01647.x

Charness, N., Yoon, J. S., Souders, D., Stothart, C., & Yehnert, C. (2018). Predictors of Attitudes towards Autonomous Vehicles: The Roles of Age, Gender, Prior Knowledge, and Personality. *Frontiers in Psychology, 9*, 2589.

Choi, J. K., & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction, 31*(10), 692-702. doi:10.1080/10447318.2015.1070549

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 ed.). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Corsi, P. (1972). *Memory and the medial temporal region of the brain.* (PhD). McGill University, Montreal, QB.

Costa Jr, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653-665. doi:10.1016/0191-8869(92)90236-I

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, *4*(1), 5-13. doi:10.1037/1040-3590.4.1.5

Crowe, S. F. (1998). The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making Test. *Journal of Clinical Psychology*, *54*(5), 585-591. doi:10.1002/(SICI)1097-4679

Das, S., Sekar, A., Chen, R., Kim, H. C., Wallington, T. J., & Williams, E. (2017). Impacts of Autonomous Vehicles on Consumers Time-Use Patterns. *Challenges, 8*(2), 1-15. doi:10.3390/challe8020032

Deary, I. J., Corley, J., Gow, A. J., Harris, S. E., Houlihan, L. M., Marioni, R. E., & Starr, J. M. (2009). Age-associated cognitive decline. *British Medical Bulletin*, *92*(1), 135-152. doi:10.1093/bmb/ldp033

Devaraj, S., Easley, R. F., & Crant, J. M. (2008). Research note—how does personality matter? Relating the five-factor model to technology acceptance and use. *Information Systems Research, 19*(1), 93-105. doi:10.1287/isre.1070.0153

Dillon, A. (2001). User Acceptance of Information Technology. In W. Karwowski (Ed.), *Encyclopedia of Human Factors and Ergonomics*. London: Taylor and Francis.

Dinh, H. Q., Walker, N., Hodges, L. F., Song, C., & Kobayashi, A. (1999). *Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments.* Paper presented at the IEEE Virtual Reality, Houston, TX.

Dotzauer, M., De Waard, D., Caljouw, S. R., Pöhler, G., & Brouwer, W. H. (2015). Behavioral adaptation of young and older drivers to an intersection crossing advisory system. *Accident Analysis & Prevention, 74*, 24-32.

Ekman, F., Johansson, M., Bligård, L.-O., Karlsson, M., & Strömberg, H. (2019). Exploring automated vehicle driving styles as a source of trust information. *Transportation Research Part F: Traffic Psychology and Behaviour, 65*, 268-279. doi:10.1016/j.trf.2019.07.026

Ekman, F., Johansson, M., & Sochor, J. (2016). *Creating appropriate trust for autonomous vehicle systems: A framework for HMI Design*. Paper presented at the Transportation Research Board 95th Annual Meeting (No. 16-3268), Washington, D.C.

Fagnant, D. J., & Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, *77*, 167-181. doi:10.1016/j.tra.2015.04.003

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. doi:10.3758/brm.41.4.1149

Ferati, M., Murano, P., & Giannoumis, G. A. (2017, July). *Universal Design of User Interfaces in Self-driving Cars.* Paper presented at the International Conference on Applied Human Factors and Ergonomics, Los Angeles, CA.

Field, A. (2009). *Discovering Statistics Using SPSS*: SAGE Publications.

Fisk, A. D., Czaja, S. J., Rogers, W. A., Charness, N., & Sharit, J. (2009). *Designing for older adults: Principles and creative human factors approaches*. Florida, FL: CRC press.

Freedman, V. A., Martin, L. G., & Schoeni, R. F. (2002). Recent trends in disability and functioning among older adults in the United States: a systematic review. *JAMA*, *288*(24), 3137-3146. doi:10.1001/jama.288.24.3137

Gkartzonikas, C., & Gkritza, K. (2019). What have we learned? A review of stated preference and choice studies on autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *98*, 323-337. doi:10.1016/j.trc.2018.12.003

Glisky, E. L. (2007). Changes in Cognitive Function in Human Aging. In D. R. Riddle (Ed.), *Brain Aging: Models, Methods, and Mechanisms*. FL: CRC Press.

González, A., Ramírez, M. P., & Viadel, V. (2015). ICT learning by older adults and their attitudes toward computer use. *Current Gerontology and Geriatrics Research, 2015*, 1-7. doi:10.1155/2015/849308

Hardman, S., Berliner, R., & Tal, G. (2019). Who will be the early adopters of automated vehicles? Insights from a survey of electric vehicle owners in the United States. *Transportation Research Part D: Transport and Environment*, *71*, 248-264. doi:10.1016/j.trd.2018.12.001

Harper, C. D., Hendrickson, C. T., Mangones, S., & Samaras, C. (2016). Estimating potential increases in travel with autonomous vehicles for the non-driving, elderly and people with travel-restrictive medical conditions. *Transportation Research Part C: Emerging Technologies*, *72*, 1-9. doi:10.1016/j.trc.2016.09.003

Harvey, J., Guo, W., & Edwards, S. (2019). Increasing mobility for older travellers through engagement with technology. *Transportation Research Part F: Traffic Psychology and Behaviour, 60*, 172-184. doi:10.1016/j.trf.2018.10.019

Hegner, S. M., Beldad, A. D., & Brunswick, G. J. (2019). In automatic we trust: investigating the impact of trust, control, personality characteristics, and extrinsic and intrinsic motivations on the acceptance of autonomous vehicles. *International Journal of Human–Computer Interaction, 35*(19), 1769-1780. doi:10.1080/10447318.2019.1572353

Holzinger A., Searle G., Nischelwitzer A. (2007). *On Some Aspects of Improving Mobile Applications for the Elderly.* In: Stephanidis C. (eds) Universal Acess in Human Computer Interaction. Coping with Diversity. UAHCI 2007. Lecture Notes in Computer Science, vol 4554. Springer, Berlin, Heidelberg. doi:10.1007/978-3-540-73279-2_103

Horberry, T., Regan, M. A., & Stevens, A. (2017). *Driver acceptance of new technology: Theory, measurement and optimisation*. CRC Press.

Hudson, J., Orviska, M., & Hunady, J. (2019). People's attitudes to autonomous vehicles. *Transportation Research Part A: Policy and Practice, 121*, 164-176. doi:10.1016/j.tra.2018.08.018

Hulse, L. M., Xie, H., & Galea, E. R. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety Science, 102*, 1-13. doi:10.1016/j.ssci.2017.10.001

IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.

International Organization for Standardization (ISO) 9241-11 (2018) Ergonomics of human-system interaction—part 11: usability: definitions and concepts. Retreived from https://www.iso.org/standard/63500.html

Jay, G. M., & Willis, S. L. (1992). Influence of direct computer experience on older adults' attitudes toward computers. *Journal of Gerontology*, *47*(4), P250-P257. doi:10.1093/geronj/47.4.P250

Jian, J. Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4*(1), 53-71. doi:10.1207/S15327566IJCE0401_04

Jokinen, J. P. P. (2015). Emotional user experience: Traits, events, and states☆. *International Journal of Human-Computer Studies, 76,* 67-77. doi:https://doi.org/10.1016/j.ijhcs.2014.12.006

Jones, M. L., Le, V. C., Ebert, S. M., Sienko, K. H., Reed, M. P., & Sayer, J. R. (2019). Motion sickness in passenger vehicles during test track operations. *Ergonomics*, *62*(10), 1357-1371. doi:10.1080/00140139.2019.1632938

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology, 3*(3), 203-220. doi:10.1207/s15327108ijap0303_3

Kyriakidis, M., de Winter, J. C., Stanton, N., Bellet, T., van Arem, B., Brookhuis, K., . . . Merat, N. (2019). A human factors perspective on automated driving. *Theoretical Issues in Ergonomics Science*, *20*(3), 223-249. doi:10.1016/j.trf.2015.04.014

Lallemand, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior, 43*, 35-48. doi:https://doi.org/10.1016/j.chb.2014.10.048

Lee, E. A.-L., Wong, K. W., & Fung, C. C. (2010). How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education, 55*(4), 1424-1442. doi:10.1016/j.compedu.2010.06.006

Li, S., Blythe, P., Guo, W., & Namdeo, A. (2019). Investigation of older drivers' requirements of the human-machine interaction in highly automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour, 62*, 546-563. doi:10.1016/j.trf.2019.02.009

Lin, J., Duh, H., Parker, D. E., Abi-Rached, H., & Furness, T. A. (2002). *Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment.* Paper presented at the IEEE Virtual Reality Orlando, FL.

Liu, P., Guo, Q., Ren, F., Wang, L., & Xu, Z. (2019). Willingness to pay for self-driving vehicles: Influences of demographic and psychological factors. *Transportation Research Part C: Emerging Technologies, 100*, 306-317.

Manawadu, U. E., Kamezaki, M., Ishikawa, M., Kawano, T., & Sugano, S. (2017). *A multimodal human-machine interface enabling situation-Adaptive control inputs for highly automated vehicles.* Paper presented at the 2017 IEEE Intelligent Vehicles Symposium (IV).

Maraj, C. S., Badillo-Urquiola, K. A., Martinez, S. G., Stevens, J. A., & Maxwell, D. B. (2017, 2017//). *Exploring the Impact of Simulator Sickness on the Virtual World Experience.* Paper presented at the Advances in Human Factors, Business Management, Training and Education, Cham.

Martens, M. H., & Jenssen, G. D. (2012). Behavioral adaptation and acceptance. In A. Eskandarian (Ed.), *Handbook of Intelligent Vehicles*. London: Springer.

Mcknight, D. H., Carter, M., Thatcher, J. B., & Clay, P. F. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems (TMIS), 2*(2), 1-24. doi:10.1145/1985347.1985353

Merat, N., Madigan, R., & Nordhoff, S. (2017). Human factors, user requirements, and user acceptance of ride-sharing in automated vehicles. Retreived from file:///C:/Users/ada.000/Downloads/human-factors-ride-sharing-automated-vehicles_0.pdf

Michel, J.-P., Beattie, L. B., Martin, F. C., & Walston, J. D. (2018). *Oxford Textbook of Geriatric Medicine* (3rd ed.). Oxford, UK: Oxford University Press.

Milleville-Pennel, I., & Charron, C. (2015). Do mental workload and presence experienced when driving a real car predispose drivers to simulator sickness? An exploratory study. *Accident Analysis & Prevention, 74*, 192-202. doi:https://doi.org/10.1016/j.aap.2014.10.021

Mirnig, A. G., Wintersberger, P., Sutter, C., & Ziegler, J. (2016, October). *A framework for analyzing and calibrating trust in automated vehicles.* Paper presented at the Proceedings of the 8th In-ternational Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct Ann Arbor, MI.

MIT Technology Reviews. (2017). Autonomous Vehicles: Are You Ready for the New Ride? Retrieved from https://www.technologyreview.com/s/609450/autonomous-vehicles-are-you-ready-for-the-new-ride/

Molnar, L. J., Ryan, L. H., Pradhan, A. K., Eby, D. W., Louis, R. M. S., & Zakrajsek, J. S. (2018). Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 58*, 319-328. doi:10.1016/j.trf.2018.06.004

Morgan, P. L., Caleb-Solly, P., Voinescu, A., & Williams, C. (2016). *Literature Review: Human-Machine Interface. Project report.* Retrieved from http://flourishmobility.com/storage/app/media/publication/Flourish%203.1.4b%20HMI%20Literature%20Review%20Short%201st%20March%202017.pdf

Morgan, P. L., Voinescu, A., Alford, C, & Caleb-Solly, Praminda. (2018). Exploring the usability of a connected autonomous vehicle human machine interface designed for older adults. Presented at: *AHFE 2018: International Conference on Applied Human Factors and Ergonomics*, Orlando, FL, USA, 21-25 July 2018. *AHFE 2018: Advances in Human Aspects of Transportation.* Advances in Intelligent Systems and Computing Springer Verlag, 591-603.

Morgan, P. L. Voinescu, A., Williams, C., Caleb-Solly, P., Alford, C., Shergold, I., Parkhurst, G., & Pipe, A. (2017). An emerging framework to inform effective design of human-machine interfaces for older adults using connected autonomous vehicles. *In: Stanton, N. ed. Advances in Human Aspects of Transportation, Vol. 597. Advances in Intelligent Systems and Computing.* Springer Verlag (Germany), Los Angeles, US. 325-334.

Mounce, R., & Nelson, J. D. (2019). On the potential for one-way electric vehicle car-sharing in future mobility systems. *Transportation Research Part A: Policy and Practice, 120*, 17-30. doi:10.1016/j.tra.2018.12.003

Musselwhite, C. (2019). Older people's mobility, new transport technologies and user-centred innovation. In *Towards User-Centric Transport in Europe* (pp. 87-103): Springer.

Naujoks, F., Wiedemann, K., Schömig, N., Hergeth, S., & Keinath, A. (2019). Towards guidelines and verification methods for automated vehicle HMIs. *Transportation Research Part F: Traffic Psychology and Behaviour, 60*, 121-136. doi:10.1016/j.trf.2018.10.012

Nikitas, A., Avineri, E., & Parkhurst, G. (2018). Understanding the public acceptability of road pricing and the roles of older age, social norms, pro-social values and trust for urban policy-making: The case of Bristol. *Cities, 79*, 78-91. doi:10.1016/j.cities.2018.02.024

Nikitas, A., Njoya, E. T., & Dani, S. (2019). Examining the myths of connected and autonomous vehicles: analysing the pathway to a driverless mobility paradigm. *International Journal of Automotive Technology and Management, 19*(1-2), 10-30. doi:10.1504/IJATM.2019.098513

Nordhoff, S., van Arem, B., & Happee, R. (2016). Conceptual model to explain, predict, and improve user acceptance of driverless podlike vehicles. *Transportation Research Record, 2602*(1), 60-67. doi:10.3141/2602-08

North, M. M., & North, S. M. (2016). A Comparative Study of Sense of Presence of Traditional Virtual Reality and Immersive Environments. *2016, 20*. doi:10.3127/ajis.v20i0.1168

Orphanides, A. K., & Nam, C. S. (2017). Touchscreen interfaces in context: a systematic review of research into touchscreens across settings, populations, and implementations. *Applied Ergonomics, 61*, 116-143. doi:10.1016/j.apergo.2017.01.013

Payre, W., Cestac, J., & Delhomme, P. (2014). Intention to use a fully automated car: Attitudes and a priori acceptability. *Transportation Research Part F: Traffic Psychology and Behaviour, 27*, 252-263. doi:https://doi.org/10.1016/j.trf.2014.04.009

Pettersson, I., & Karlsson, I. M. (2015). Setting the stage for autonomous cars: a pilot study of future autonomous driving experiences. *IET intelligent transport systems, 9*(7), 694-701. doi:10.1049/iet-its.2014.0168

Price, M., Mehta, N., Tone, E. B., & Anderson, P. L. (2011). Does engagement with exposure yield better outcomes? Components of presence as a predictor of treatment response for virtual reality exposure therapy for social phobia. *Journal of Anxiety Disorders, 25*(6), 763-770. doi:10.1016/j.janxdis.2011.03.004

Rebelo, F., Noriega, P., Duarte, E., & Soares, M. (2012). Using virtual reality to assess user experience. *Human Factors, 54*(6), 964-982. doi:10.1177/0018720812465006

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8*(3), 271-276. doi:10.2466/pms.1958.8.3.271

Riener, A., Boll, S., & Kun, A. L. (2016). *Automotive user interfaces in the age of automation (Dagstuhl Seminar 16262).* Paper presented at the Dagstuhl reports. Retreived from https://drops.dagstuhl.de/opus/volltexte/2016/6758/

Ritchie, O. T., Watson, D. G., Griffiths, N., Misyak, J., Chater, N., Xu, Z., & Mouzakitis, A. (2019). How should autonomous vehicles overtake other drivers? *Transportation Research Part F: Traffic Psychology and Behaviour, 66*, 406-418. doi:10.1016/j.trf.2019.09.016

Rizzo, A. S., & Koenig, S. T. (2017). Is clinical virtual reality ready for primetime? *Neuropsychology, 31*(8), 877-899. doi:10.1037/neu0000405

SAE International. (2016). U.S. Department of transportation's new policy on automated vehicles adopts SAE International's levels of automation for defining driving automation in on-road motor vehicles. Retrieved from https://www.sae.org/news/3544

Sanbonmatsu, D. M., Strayer, D. L., Yu, Z., Biondi, F., & Cooper, J. M. (2018). Cognitive underpinnings of beliefs and confidence in beliefs about fully automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour, 55*, 114-122. doi:10.1016/j.trf.2018.02.029

Sârbescu, P., & Neguţ, A. (2013). Psychometric properties of the Romanian version of the Zuckerman-Kuhlman Personality Questionnaire. *European Journal of Psychological Assessment, 29*(4), 241-252. doi:10.1027/1015-5759/a000152

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors, 58*(3), 377-400. doi:10.1177/0018720816634228

Schoettle, B., & Sivak, M. (2014). *A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia*. Retrieved from https://deepblue.lib.umich.edu/handle/2027.42/108384

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition and Emotion*, *14*(4), 433-440. doi:10.1080/026999300402745

Shergold, I. (2018). Findings from workshops held with older people considering participating in connected autonomous vehicle trials. Retreived from https://uwe-repository.worktribe.com/output/868759/findings-from-workshops-held-with-older-people-considering-participating-in-connected-autonomous-vehicle-trials

Shergold, I., Lyons, G., & Hubers, C. (2015). Future mobility in an ageing society–Where are we heading? *Journal of Transport & Health, 2*(1), 86-94. doi:10.1016/j.jth.2014.10.005

Shladaver, S. E. (2016). The truth about "self-driving" cars. *Scientific American, 314*(6), 52-57. Retreived from https://www.scientificamerican.com/article/the-truth-about-ldquo-self-driving-rdquo-cars/

Siren, A., & Haustein, S. (2015). Driving licences and medical screening in old age: re-view of literature and European licensing policies. *Journal of Transport & Health, 2*(1), 68-78. doi:10.1016/j.jth.2014.09.003

Souders, D., & Charness, N. (2016, July). *Challenges of older drivers' adoption of advanced driver assistance systems and autonomous vehicles.* Paper presented at the International Conference on Human Aspects of IT for the Aged Population, Toronto, ON, Canada.

Spurlock, C. A., Sears, J., Wong-Parodi, G., Walker, V., Jin, L., Taylor, M., ... & Todd, A. (2019). Describing the users: Understanding adoption of and interest in shared, electrified, and automated transportation in the San Francisco Bay Area. *Transportation Research Part D: Transport and Environment*, *71*, 283-301. doi:10.1016/j.trd.2019.01.014

Strauch, C., Mühl, K., Patro, K., Grabmaier, C., Reithinger, S., Baumann, M., & Huckauf, A. (2019). Real autonomous driving from a passenger's perspective: Two experimental investigations using gaze behaviour and trust ratings in field and simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, *66*, 15-28. doi:10.1016/j.trf.2019.08.013

Svendsen, G. B., Johnsen, J. A. K., Almås-Sørensen, L., & Vittersø, J. (2013). Personality and technology acceptance: the influence of personality factors on the core constructs of the Technology Acceptance Model. *Behaviour & Information Technology*, *32*(4), 323-334. doi:10.1080/0144929X.2011.553740

Swan, G., Shahin, M., Albert, J., Herrmann, J., & Bowers, A. R. (2019). The effects of simulated acuity and contrast sensitivity impairments on detection of pedestrian hazards in a driving simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, *64*, 213-226. doi:10.1016/j.trf.2019.05.003

Talebpour, A., & Mahmassani, H. S. (2016). Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C: Emerging Technologies, 71*, 143-163. doi:10.1016/j.trc.2016.07.007

Tcha-Tokey, K., Loup-Escande, E., Christmann, O., & Richir, S. (2016). *A questionnaire to measure the user experience in immersive virtual environments*. Paper presented at the Proceedings of the 2016 Virtual Reality International Conference, Laval, France.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, *28*, 127-154. doi:10.1016/0749-596X(89)90040-5

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*(3), 498-505. doi:10.3758/BF03192720

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425-478. doi:10.2307/30036540

Voinescu, A., Morgan, P. L., Alford, C., & Caleb-Solly, P. (2018). Investigating older adults' preferences for functions within a human-machine interface designed for fully autonomous vehicles. In *International Conference on Human Aspects of IT for the Aged Population* (pp. 445-462). Springer, Cham.

Wang, Z., Vang, M., Lookadoo, K., Tchernev, J. M., & Cooper, C. (2014). Engaging high-sensation seekers: The dynamic interplay of sensation seeking, message visual-auditory complexity and arousing content. *Journal of Communication, 65*(1), 101-124. doi:10.1111/jcom.12136

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070. doi:10.1037//0022-3514.54.6.1063

Webb, J., Wilson, C., & Kularatne, T. (2019). Will people accept shared autonomous electric vehicles? A survey before and after receipt of the costs and benefits. *Economic Analysis and Policy, 61*, 118-135. doi:10.1016/j.eap.2018.12.004

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225-240. doi:10.1162/105474698565686

Wu, J., Liao, H., Wang, J.-W., & Chen, T. (2019). The role of environmental concern in the public acceptance of autonomous electric vehicles: A survey from China. *Transportation Research Part F: Traffic Psychology and Behaviour, 60*, 37-46. doi:10.1016/j.trf.2018.09.029

Yi, Z., Smart, J., & Shirk, M. (2018). Energy impact evaluation for eco-routing and charging of autonomous electric vehicle fleet: Ambient temperature consideration. *Transportation Research Part C: Emerging Technologies, 89,* 344-363. doi:10.1016/j.trc.2018.02.018

Zhang, T., Tao, D., Qu, X., Zhang, X., Lin, R., & Zhang, W. (2019). The roles of initial trust and perceived risk in public's acceptance of automated vehicles. *Transportation Research Part C: Emerging Technologies, 98*, 207-220. doi:10.1016/j.trc.2018.11.018

Table 1. *Current AV/CAV HMI design principles*

Author/s	Population	Type of study	Automation level	Summary of findings
Petterson & Karlsson, 2015	Adult population	Participatory design, mixed methods	Level 3	Users expect: increasing social capabilities and novel in-vehicle technologies and to accommodate work and leisure activities; route and driving information should be available at all times; prefer voice functionality; Various activities during journey: relaxation, working, sleeping, reading, socializing, video entertainment, games and social media.
Naujoks et al., 2019	Adult population	Literature review	Levels 3 to 4	Importance of up-dating the current HMI principles to accommodate the challenges AVs pose. Key features of HMIs concern: handover and hand-back control (e.g., avoid unintentional activation and deactivation); system mode (always display system mode, and any changes); visual messages, auditory and tactile input that require immediate action should be delivered differently than non-critical action; high priority input should be delivered multimodal; system failure should inform the driver about the cause/location of the problem and provide operator steps.
Ferati et al., 2017	Adult population	Literature review	Levels 3 to 5	Multimodal HMIs might both compensate and attract user attention, when they are engaged in other tasks (e.g. reading a book, checking emails)
Das et al., 2017	Adult population	Survey-based study	Level 5	AV HMI features that are likely to be used by e.g., commuters are: internet connectivity; accessible computers that are comfortable to use; displays that support media watching; comfortable seats that can be used for sleeping.
Ekman et al., 2019	Adult population	Simulator-based study	Level 5	Users prefer a "defensive" driving style as this is more predictable in opposition to "aggressive" driving style.
Morgan et al., 2017	Older adults	Systematic literature review	Level 5	Reviewed and synthetized CAV HMI design principles for older adults based on four areas:

				usability (e.g., reduce clutter and use simple displays; minimise the number and frequency of distractor stimuli; organise information in a natural and consistent way); accessibility (e.g., use simple, minimal and intuitive steps in order to perform tasks; use large screens with large buttons; make information clearly visible using size, colour and contrast features); functionality (e.g., provide information on vehicle speed and journey time; remove complex interfaces and integrate traditional navigation methods); adaptability (e.g., provide the option to zoom-in and out).
Li et al., 2019	Older adults		Level 4	Older participants prefer simple and safe AV HMIs that are similar to traditional HMIs. Positive towards the possibility of engaging in a range of non-driving related tasks while the vehicle was in full AV mode. Prefer AV HMIs that always display the system mode and any changes (e.g., full AV driving mode). Preference towards a balance between messages displayed by the HMI and avoidance of clutter.

Table 2. *Personality and AV and/or CAV preferences*

Author/s	Population studied	Type of study	Summary of findings
Charness et al., 2018	Adult and older adult population; no exclusion criteria was specified	Survey-based study	Conscientiousness is a positive predictor of concern with AVs and negative predictor of eagerness to adopt them. Emotional stability and openness were significant positive predictors of eagerness to adopt AVs. Extraversion and openness were significant predictors of willingness to relinquish driving control.
Bennett et al., 2019	People with intellectual disabilities and mental health problems	Survey-based study	Prior knowledge of AVs and intensity of disability positively predict the willingness to travel in an AV. Age, gender, income, locus of control, and anxiety were not significant predictors.
Payre et al, 2014	Adult and older adult population; no exclusion criteria was specified	Survey-based study	Positive correlations between the level of sensation seeking and the intention of using AVs.
Kyriakidis et al., 2015	Adult and older adult population; no exclusion criteria was specified	Survey-based study	High neuroticism was associated with concerns about data transmitting. Positive correlation between agreeableness and data transmitting.
Voinescu et al., 2018	Older adults and people with sensory and/or cognitive impairments	Survey-based study	Neuroticism anxiety personality trait correlated negatively with the preference for the following AV HMI functions: television, news, and weather search. Activity, a personality trait that describes the tendency to work hard and be involved in many activities, correlated positively with the preference for weather search. Aggressive hostility trait correlated negatively with the likelihood of using a television function. Sociability trait correlated negatively with view planned journey function (the AV HMI function that allows the user to search for planned journeys, similar to past journeys in current SatNavs).

			Impulsive sensation seeking did not correlate with the likelihood of using any function.
Spurlock et al., 2019	Adult population; no exclusion criteria was specified	Survey-based study	Risk loving individuals are less likely to adopt AVs. Agreeableness does not predict adoption intention of AVs.
Hegner et al., 2019	Adult population; no exclusion criteria was specified	Survey-based study	Personal innovativeness positively influenced the adoption intention of AVs.

Note. The methodology described in survey-based studies did not include exposure of participants in AV simulators.

Table 3. *Sample demographics*

Variable	Summary statistics: Mean, SD, %
Age	70.2 *(4.46)*
Gender	
Male	72.0 %
Female	28.0%
Highest qualification	
O Level, GCSE, or equivalent or less	3.2%
A Level, AS level, or equivalent	6.5%
Further Education or vocational training (including HND)	25.8%
First Degree (e.g. BSC, BA)	19.4%
Higher Degree (e.g. MSc, MA, PhD)	9.7%
Postgraduate training (separate to a PG degree)	16.1%
Work status	
Full time	4.0%
Part time	12.0%
Retired	84.0%
Other	0.0%
Marital status	
Single	16.0%
Married	60.0%
De facto/have a partner	8.0%
Divorced	8.0%
Widowed	8.0%
Driving license	
Yes	92.0%
No	8.0%
Average miles driven in the last 12 months	
Less than 1000	24.0%
1000 to 2400	4.0%
2500 to 4900	16.0%
5000 to 7400	28.0%

7500 to 9900	4.0%
10000 to 14900	16.0%
15000 to 19900	0%
Over 20000	0%
N/a	8.0%
Speeding offences in the last 12 months	
Yes	20%
No	72%
N/a	8.0%
Penalty points on the driving licence	
Yes	18%
No	76%
N/a	8%

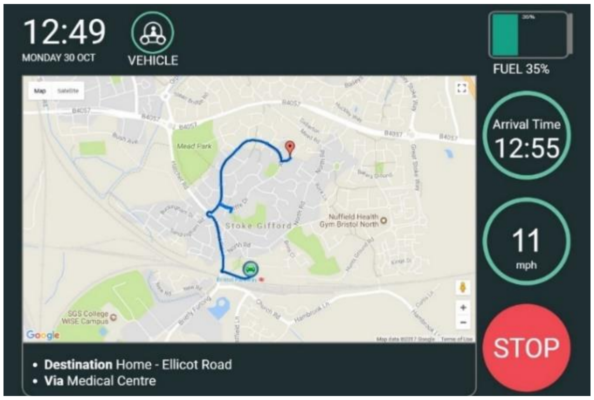

Figure 1. *Main dashboard of the HMI used*

Figure 2. *Vehicle status button*

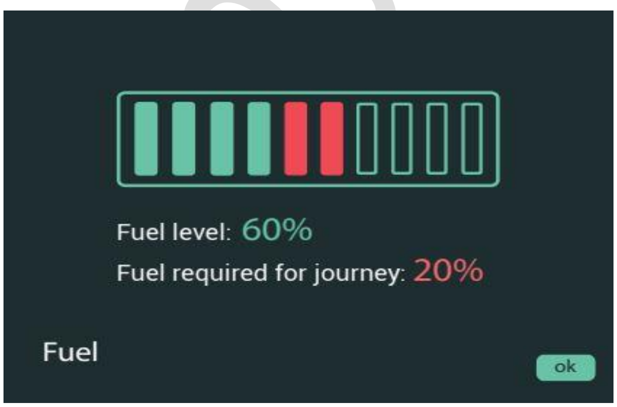

Figure 3. *Fuel status button*

Figure 4. *Stop button*

Figure 5. *Exterior of the Lutz pod*

Figure 6. *Interior of the Lutz pod*

Figure 7. *Simulator integration/virtual environment used in our study*

Appendix A. Supplementary material

Supplementary Table 1. *HMI features*

Button	Button status (active vs inactive)	Features
Date and Time	Passive button	Displays date and hour in real time (see Figure 1)
Vehicle Status	Interactive button	If pressed a pop-up button opened and showed the status of the vehicle (see Figure 2)
Fuel	Interactive button	If pressed a pop-up button with fuel status opened (see Figure 3)
Arrival Time	Passive button	Displays arrival time in real time (see Figure 1)
Speed	Passive button	Displays speed in miles per hour in real time (see Figure 1)
Emergency Stop	Interactive button	If pressed a pop-up button opened with 2 options: No or Yes (see Figure 4)
Navigation Map	Interactive button	Google-based map with position of the vehicle tracked, supports zoom in and zoom out

Supplementary Table 2. *Full description of the measures used*

Psychological construct	Measure	Platform	Scale/task description	Outcomes
Working memory capacity task and multi-tasking abilities	Ospan (Unsworth et al., 2005)	Computerized task implemented with Pebl Version 2.0 Beta 4, Mueller & Piper, 2014)	The Opsan task contained two alternate sub-tasks: 1) remember a serious of two to five letters, and 2) solve math problems (distractor task). The set size of the letters and math problems could not be anticipated by the participants because the trials were quasi-randomized. The performance on the letter tasks was calculated by giving equal points to the set size, but only if all the letters from that set were recalled correctly in serial order. This resulted in an absolute span score. During the math task, the accuracy of participants' responses was tracked and participants received feedback for it. The feedback was used to maintain performance accuracy ≥85% and to keep participants engaged with the task.	We used two outcomes: 1) total correct letters recalled and 2) absolute span score. Higher scores reflect better performance.
Executive function	Trail Making Test Form A & B (TMT A & B, Reitan, 1958)	Computerized task which ran on a LearnPad Android device with Pen Six Screener (PenScreenSix Cognitive Testing Software v2.0 for Android, 2014)	TMT has 2 forms: A and B. For A, the participant has to connect, as quickly as possible, 25 encircled numbers in ascending order. For B, the participant tries to connect, as quickly as possible, numbers and letters in ascending order.	The outcomes are: (1) time to complete the task, (2) time to complete the task minus the first 2 responses, and (3) number of incorrect responses. Lower scores represent better performance.
Spatial working memory span	Corsi Blocks Test (Corsi, 1972)	Computerized task implemented with Pebl	The task consisted of sequences of blocks displayed irregularly on the desktop screen. The participant was	The following outcomes were used: 1) block span, 2) total correct trials, and 3) the

		(Version 2.0 Beta 4, Mueller & Piper, 2014)	instructed to tap the blocks on the screen as they light up one by one in a random order. The task includes 12 trials that start with 2 letters and continue up to 7 letters (e.g. 2 trials with 2 letters, 2 trials with 3 letters, 2 trials with 4 letters, etc.).	product. Higher scores represent better performance.
Personality	Zuckerman-Kuhlman Personality Questionnaire (ZKPQ-50-CC, shortened form (Aluja et al., 2006)	Paper-and-pencil	The questionnaire is based on the Alternative Five Factor Model (Aluja et al., 2002; 2006; Zuckerman, 2014) and measures five personality traits that arguably best describe human behavior: impulsive sensation seeking (ImpSS) (e.g. lack of planning, tendency to act quickly on impulse, risk taking, novelty seeking), aggression hostility (Agg-Host) (e.g. antisocial behavior, vengefulness, quick temper), sociability (Sy) (e.g. having many friends, enjoying large parties, intolerance for social isolation), activity (Act) (e.g. need for general activity, impatience, preferences for challenging and hard work), and neuroticism anxiety (N-Anx) (e.g. emotional upset, worry, tension, obsessive indecision).	The ZKPQ-50-CC has 50 true-false items, and the total score for each trait is computed by giving a 0 (for a No answers) or 1 (for Yes answers), noting that the questionnaire has reversed items. Original scoring procedure can be found in the original paper (Zuckerman, 2014; Aluja et al., 2002). Higher scores reflect increased personality traits (e.g. someone scoring high on sociability is more sociable than someone scoring low in this trait). For impulsive sensation seeking internal consistency was questionable (Cronbach's α = .65), for aggression hostility it was poor (Cronbach's α = .56), for sociability it was questionable (Cronbach's α = .69), for activity it was good (Cronbach's α = .80), and for

				neuroticism anxiety it was good (Cronbach's α = .84). Values ranging from .71 for sociability to .79 for neuroticism anxiety are reported by Aluja et al. (2007), and from .60 for aggression hostility to .82 for neuroticism anxiety on a German sample and from .72 for aggression hostility to .80 for neuroticism anxiety on an American sample (Aluja et al., 2006).
Trust in technology trait	General Trust in Technology Scale (GTS, Mcknight et al., 2011)	Paper-and-pencil	It measures trust trait and contains 7 questions that measure people's trust in technology (e.g. I believe that most technologies are effective at what they are designed to do; I think most technologies enable me to do what I need to do). Responses are recorded using a Likert scale with 1-7 gradations (Mcknight et al., 2011).	Higher scores represent increased trust in technology. The scale has reversed items. Higher scores represent increased trust in technology. Internal consistency was good (Cronbach's α = .80).
Attitudes towards computers	Attitudes Towards Computers Questionnaire (ATCQ, Jay & Willis, 1992)	Paper-and-pencil	For example, how people relate to computers and whether they are willing to use them for personal or professional reasons. The scale contains 32 items (e.g. I feel comfortable with computers; Computers are making the jobs done by humans less important) with response options on a 5-point Likert scale format.	It is a multidimensional scale with 7 scales: Comfort, Efficacy, Gender Equality, Control, Dehumanization, Interest and Utility. Lowers scores reflect more negative attitudes towards computers. Internal consistency was as follows: unacceptable for Comfort (Cronbach's α = .34),

				questionable for Efficacy (Cronbach's α = .62), unacceptable for Gender Equality (Cronbach's α = .16), acceptable for Control (Cronbach's α = .71) and for Dehumanization (Cronbach's α = .74), poor for Interest (Cronbach's α = .54), and unacceptable for Utility (Cronbach's α = .35). As a result, we chose to use only Efficacy, Control, and Dehumanization scales.
Simulator sickness	Simulator Sickness Questionnaire (SSQ, Kennedy et al., 1993)	Paper-and-pencil	The questionnaire consists of a checklist of sixteen symptoms that usually appear if/when experiencing simulator sickness (e.g. dizziness, nausea) described within the questionnaire.	Participants are instructed to rate on a scale of 0-3 the severity of symptoms, if any (e.g., general discomfort, blurred vision, dizziness with eyes open, nausea). A total score above 20 indicates that participants experience simulator sickness (Stanney, Kennedy, & Drexler, 1997) Higher scores represent increased simulator sickness. Internal consistency was acceptable (Cronbach's α = .72).
Presence	Presence Questionnaire (PQ, Witmer & Singer, 1998)	Paper-and-pencil	It contains 22 items that measure the level of subjective immersion and presence in a virtual environment defined by a sense of being present	Participants have to rate their responses on a 7-point Likert scale. Lower scores indicate a diminished sense of presence

			("there") in the virtual environment (e.g. How much did the visual display quality interfere or distract you from performing assigned tasks or required activities?).	in the virtual environment. Internal consistency was questionable (Cronbach's α = .66)
Mood	Positive and Negative Affect Schedule (PANAS, Watson et. al., 1988)	Paper-and-pencil	Mood was measured with It is one of the most widely used measure of emotions and contains 20 items that describe positive and negative emotions (e.g. excited, guilty).	Participants have to rate the extent to which they currently felt emotions (e.g. distressed, alert) on a 5 point Likert-type scale. Lower scores represent reduced levels of that particular emotion.
Trust in automation state	Checklist for Trust between People and Automation (CTBPA, Jian et al., 2000)	Paper-and-pencil	Trust in automation state is measured with 12 questions on aspects such as dependability and reliability of the system, suspicion, and confidence.	Higher scores imply increased system dependability and ratability. Internal consistency was good (Cronbach's α = .88).
System usability	System Usability Scale (SUS, Brooke, 1986)		Measure of perceived usability and user experience. It contains 10 items (e.g. I found the various functions in this system were well integrated).	Participants have to rate on a 5-point Likert scale ranging from 1 "totally disagree" to 5 "totally agree". Higher scores reflect better usability of the AV HMI. Internal consistency was questionable (Cronbach's α = .68).

Supplementary Table 3. *Correlation coefficients (Pearson r) for user experience (as measured with the System Usability Scale) and cognitive processes (Corsi, Ospan and Trails)*

Measure

	Corsi Block Span	Corsi Total Correct	Corsi Product	Ospan Total Correct Letters	Span score/ Absolute score	Trails A CT25	Trails A CT23	Trails A NI	Trails B CT25	Trails B CT23	Trails B NI
SUS	.47**	.52**	.44*	.59**	.53**	-.08 *ns*	-.04 *ns*	-.04 *ns*	-.48**	-.45*	-.19 *ns*

Note. ** $p < .01$, * $p < .05$, *ns*- nonsignificant. SUS = System Usability Scale (Brooke, 1986); Corsi Block Span = Corsi Blocks Test (Corsi, 1972); Corsi Total Correct = Corsi Blocks Test (Corsi, 1972); Corsi Product = Corsi Blocks Test (Corsi, 1972); Ospan Total Correct Letters = Operation Span/Ospan (Turner & Engle, 1989; Unsworth et al., 2005); Span Score/Absolute Score = Operation Span/Ospan (Turner & Engle, 1989; Unsworth et al., 2005); Trails A CT25 = Trail Making Test Part A (Reitan, 1958); Trails A CT23 = Trail Making Test Part A (Reitan, 1958); Trails A NI = Trail Making Test Part A (Reitan, 1958); Trails B CT25 = Trail Making Test Part B (Reitan, 1958); Trails B CT23 = Trail Making Test Part B (Reitan, 1958); Trails B NI = Trail Making Test Part B (Reitan, 1958)

Supplementary Table 4. *Correlation coefficients (Pearson r) for user experience (as measured with the System Usability Scale) and personality (ZKPQ), attitudes toward computers (ATCQ) and trait trust in technology (GTS)*

Measure

	ZKPQ Act	ZKPQ Agg-Host	Sy	N-Anx	ImpSS	ATCQ Efficacy	ATCQ Control	ATCQ Dehumanization	GTS
SUS	-.12 *ns*	-.00 *ns*	-.02	-.37*	.32 *ns*	-.06 ns	.01ns	-.31 ns	.47*

Note. * $p < .05$, *ns*- nonsignificant. SUS = System Usability Scale (Brooke, 1986); ZKPQ Act = Activity, Zuckerman-Kuhlman Personality Questionnaire shortened form, Alternative Five Factor Model (Aluja et al., 2006); ZKPQ Agg-Host = Aggression hostility, Zuckerman-Kuhlman Personality Questionnaire shortened form, Alternative Five Factor Model (Aluja et al., 2006); ZKPQ Sy = Sociability, Zuckerman-Kuhlman Personality Questionnaire shortened form, Alternative Five Factor Model (Aluja et al., 2006); ZKPQ N-AnxN = Neuroticism anxiety, Zuckerman-Kuhlman Personality Questionnaire shortened form, Alternative Five Factor Model (Aluja et al., 2006); ZKPQ ImpSS = Impulsive Sensation Seeking, Zuckerman-Kuhlman Personality Questionnaire shortened form, Alternative Five Factor Model (Aluja et al., 2006); ATCQ Efficacy = Efficacy, Attitudes Toward Computers Questionnaire (Jay & Willis, 1992); ATCQ Control= Control, Attitudes Toward Computers Questionnaire (Jay & Willis, 1992); ATCQ Dehumanization = Dehumanization, Attitudes Toward Computers Questionnaire (Jay & Willis, 1992); GTS = General Trust Scale (McKnight et al., 2011).

Supplementary Table 5. *Correlation coefficients (Pearson r) for user experience (as measured with the System Usability Scale), presence (PQ), simulator sickness (SSQ) and state trust in automation (CTBPA)*

Measure

	PQ	SSQ	CTBPA
SUS	.33 *ns*	-.37*	.05 *ns*

Note. * $p < .05$, *ns*- nonsignificant. SUS = System Usability Scale (Brooke, 1986); PQ = Presence Questionnaire (Witmer & Singer, 1998); SSQ = Simulator Sickness Questionnaire (Kennedy et al., 1993); CTBPA = Checklist for Trust between People and Automation (Jian et al., 2000).

Supplementary Table 6. *Correlation coefficients (Pearson r) for user experience (as measured with the System Usability Scale), positive and negative emotions (PANAS)*

Measure

	Positive emotions				Negative emotions			
	J1	J2	J3	J4	J1	J2	J3	J4
SUS	.59**	.38*	.59**	.59**	-.28 *ns*	-.22 *ns*	-.53**	-.58**

Note. ** $p < .01$, * $p < .05$, *ns*- nonsignificant. SUS = System Usability Scale (Brooke, 1986); Positive emotions = PANAS Positive emotions, Watson et al., 1988; Negative emotions = PANAS Negative emotions, Watson et al., 1988; J1 = 1st journey in the CAV simulator; J2 = 2nd journey in the CAV simulator; J3 = 3rd journey in the CAV simulator; J4 = 4th journey in the CAV simulator.

Design and Data Analysis

For comparisons between journey order (counterbalanced) and in case of differences due to positive and negative emotions (PANAS), a repeated measures design was employed and a repeated measure analysis of variance (ANOVA) was used to test the assumption. A paired-sample *t* test was used for comparisons between: pre-test simulator sickness and post-test simulator sickness. Independent sample *t* tests were used to compare older and younger older adults (cut-off 70 years old), plus males and females on usability (SUS).

Results

Results from a paired-sample *t*-test showed that at post-test, the level of reported simulator sickness (M = 2.88; SD = 2.72) was significantly higher than at pre-test (M = 0.68; SD = 1.46), $t(30) = -3.79$, $p < .01$. However, none of the older participants reported moderate-severe simulator sickness symptoms (resulting in voluntary or encouraged withdrawal from the study). The total score on SSQ did not exceed the score of 9, and moderate simulator sickness accounts for a score larger than 15 (Stanney, Kennedy, & Drexler, 1997). A repeated measures ANOVA using Pillai's trace revealed a non-significant main effect of journey order both on positive emotions, $V = 0.14$, $F(4, 22) = 1.26$, $p > .05$, and on negative emotions, $V = 0.12$, $F(4, 22) = 1.08$, $p > .05$. There was no difference between males and females on user experience of the CAV HMI, $t(23) = -0.39$, $p > .05$. Both females and males were comparable in terms of their ratings of system usability.

An independent *t*-test revealed no significant differences between younger-older participants (cut-off 70 years old) and older-older adults (71-years+) in terms of their user experience with the AV HMI, $t(23) = 0.69$, $p > .05$.