Stimulating Conversation: Engaging Drivers in Natural Language Interactions with an Autonomous Digital Driving Assistant to Counteract Passive Task-Related Fatigue

David R. Large[1*], Gary Burnett[1], Vicki Antrobus[1], Lee Skrypchuk[2]

[1] Human Factors Research Group, Faculty of Engineering, University of Nottingham, UK
[2] Jaguar Land Rover Research, International Digital Laboratory, Coventry, UK
[*]david@largebrown.co.uk

Abstract: Engaging in conversation has been shown to be an effective countermeasure to passive task-related fatigue. To investigate the effectiveness of a digital assistant to fulfil the role of conversational partner in counteracting driver fatigue, twenty participants undertook two 30-minute drives in a medium fidelity driving simulator, within a low-feature, monotonous driving environment – following a lead car at 68-mph in lane one of a UK-style motorway. All testing occurred between 13:00 and 16:30, when circadian and homeostatic influences naturally reduced participants' alertness. During one of the counterbalanced drives, participants engaged in natural language interactions with a state-of-the-art digital driving assistant, delivered using a Wizard-of-Oz approach. Results suggest that the digital assistant had a positive effect on driver alertness compared with the control condition (no assistant): there was a trend towards lower perceived sleepiness and significantly higher arousal after driving with the digital assistant. Objectively, interacting with the digital assistant improved lane keeping (lower SDLP) and attracted earlier responses to a hazard situation. There were also significant differences in eye activity when conversing with the digital assistant, revealed by fewer fixations directed towards the road centre and larger pupil diameter (suggesting increased arousal). The findings have implications for the design of future in-vehicle natural language interfaces.

1. Introduction

Driver fatigue is frequently cited as a causal or contributory factor to road traffic accidents. Analysis of accident data suggests that it is responsible for as many as 20% of all road accidents [1]. In addition, the early signs of fatigue, such as inattention, poor decision making, delayed reaction times and driver errors, are thought to have an equally profound and debilitating effect on driving performance and contribute to a much higher proportion of collisions [2].

1.1. Fatigue and Sleepiness

Fatigue is a state of reduced mental alertness that impairs performance. It is defined as a gradual and cumulative process associated with "a loss of efficiency, and a disinclination for any kind of effort" [3]. Fatigue can occur due to a numbers of factors – acting together or in isolation – including interruption of circadian rhythms (the 'body clock'), low brain activation levels and sustained attention; it is therefore particularly prevalent during night-time [4] and monotonous driving [5], and in situations where driving hours are long [2].

If uninterrupted, fatigue can naturally lead to sleep [6]. The term *sleepiness* is therefore often used interchangeably with *fatigue* [2]. Although the terms exist as lucid concepts within the literature, the concepts of fatigue and sleepiness are inextricably linked – fatigue can promote sleepiness and sleepiness can elevate feelings of fatigue [7]. Moreover, the general sensation of weariness, feelings of inhibition and impaired activity – that define *fatigue* – are symptoms that also pre-empt or signify the onset of *sleep* [2]. The terms fatigue and sleepiness therefore appear to be used inconsistently and synonymously, particularly within automotive human factors literature and, for the purposes of this paper, we will maintain this convention.

In contrast, a clear distinction can be drawn from the literature between *mental* and *physical* fatigue – mental fatigue is believed to be psychological in nature – caused by impairment of the cognitive functions of the driver – whereas physical fatigue is considered synonymous with muscle fatigue [6]. In an automotive context, where the role of a driver is nowadays largely cognitive in nature, requiring sustained vigilance, selective attention, complex decision-making and occasional automatized perceptual-motor control skills, mental fatigue is a particular concern. It has also been shown that mental fatigue, in particular, is affected by the changing levels of arousal produced by circadian rhythms. This can lead to deteriorations in performance, for example during afternoon (or night-time) driving when physiological activity is diminishing [8].

1.2. Measuring Fatigue

Methodological approaches for assessing the onset and progression of driver fatigue are abundant and extremely varied, although there remains some debate regarding the interpretation of data and reliability of many of these measures (see: [6]). It is therefore prudent to use a number of different measures during fatigue-related studies. Common metrics are based on physiological measurements (e.g. heart activity, electrodermal activity [6, 9]) – thought to be synonymous with indicators of individual physiological level of fatigue/workload; driving performance (speed maintenance, headway variability, lane keeping, e.g. [10, 11]), behavioural indicators/eye activity (blink-rate, PERCLOS, PRC [12, 13]) and subjective assessment (Epworth and Stanford Sleepiness scale [14, 15]) are also commonly employed.

In a driving context, fatigue can effect performance by increasing the frequency, amplitude and/or variability of errors, with studies demonstrating increased lane drift [10, 16, 17]), poor speed control [17] and late corrections to lane positioning [10, 11] amongst fatigued drivers. Fatigue may be quite severe before routine driving performance is noticeably affected [6] – particularly amongst professional drivers – however, decreases in physiological arousal, slowed sensorimotor functions and impaired information processing may be evident at much lesser levels of fatigue. This can retard a driver's vigilance and their ability to respond to unusual and emergency events, resulting in slower reaction times to traffic controls and hazards [18].

Heart-rate and heart-rate variability are also generally considered to be a good relative indicator of workload/fatigue, but effects can be inconsistent. For example, studies show that heart-rate increases and heart-rate variability decreases during demanding mental processing, although it has also been shown that heart-rate decreases significantly during a monotonous driving task [19].

In addition to changes in driving performance and physiological arousal, drivers exhibit certain behavioural characteristics when fatigued, such as changes to eye activity (e.g. elevated blink duration and frequency) [19]. Research has also shown that quiescence in eye movements is one of the earliest reliable signs of sleepiness [20], thus one can expect increases in the duration (and reductions in frequency) of visual fixations to accompany the onset and escalation of fatigue. In particular, fatigued drivers are likely to focus on the road centre at the expense of other driving related objects such as signs, bicyclists, scenery etc. in the forward and peripheral road scene [13]. Percent Road Centre (PRC) is thus defined as the proportion of time that a driver's eyes are focussed on the *road centre* – defined by a 20° (horizontal) x 15° (vertical) rectangular area centred around the driver's mean point of fixation [13]. PRC has been shown to be sensitive to secondary task workload [13] and is thought to be equally perspicuous in situations of driver fatigue.

Self-report techniques are also commonly employed to determine driver alertness and emotional state. These are based on the assumption that an individual's subjective experience of a particular construct reflects the objective reality of that measure. In general, two approaches exist regarding fatigue/sleepiness: determining the driver's current perception of their level of sleepiness, i.e. sleepiness as a state characteristic (e.g. Stanford Sleepiness Scale [15]), or gaining a general perception of their propensity for sleepiness as a component of daily life, i.e. sleepiness as a trait characteristic (e.g. Epworth Sleepiness Scale [14]).

1.3. Countermeasures for Fatigue

Anecdotally, drivers appear to be aware of the risks associated with fatigue, yet in practice are often reluctant or unwilling to instigate effective countermeasures (or preventive strategies) when sleep beckons – research suggests that many drivers continue to drive tired despite being aware of their fatigued state [1], rather than following advice to take a break from driving, consume caffeine and have a short nap [21]. It is likely that some drivers fail to *fully* appreciate the risks associated with fatigued driving, or justify their actions based on past experience, perceived societal norms and journey goals. Public awareness campaigns, such as 'THINK!' in the UK [22], are an effective method to improve drivers' understanding of the dangers of driving while fatigued, but still require drivers to actively engage in preventive measures. Furthermore, the low arousal state and reduced brain activation levels associated with fatigue means that drivers may not be consciously aware of the extent of their fatigued state, crucially at the time that they are most at risk, and may therefore lack the wherewithal to initiate effective proactive countermeasures. In this situation, enforcing operational countermeasures – intended to enhance alertness and performance temporarily so that operational safety and performance are maintained [23] – may be more effective.

A variety of in-vehicle devices have also been proposed by vehicle manufacturers (see: [24] for an early review). However, many of these devices only respond to the later signs of fatigue, and rely on physical indicators, such as elevated blink-rate, eye and eyelid activity, and mannerisms such as nodding and yawning. The concern is that by the time that fatigue is identified in this manner, a driver's performance is likely to already be significantly impaired. Moreover, such interventions may also encourage greater risk taking, shifting the responsibility for recognising fatigue from the driver to the technology – a fatigued driver may continue driving with the belief that the system will alert them if necessary [1].

One of the most effective operational strategies to combat the onset of fatigue, and associated performance decrements, is social interaction and conversation [23, 25, 26]. Social interaction, particularly conversation, has been shown to maintain alertness amongst pilots even during the circadian nadir (the lowest point of natural circadian fluctuations) [26]. Moreover, research has shown that in some situations (e.g. pilots during long-haul flights), the *absence* of conversation can be a predictor of *declining* physiological alertness [27]. To maintain alertness in this manner, the protagonist must be actively involved in the conversation and not just listening [23]: this therefore requires a second interlocutor, with whom to converse – a role that comfortably fits a co-pilot in aviation and is often naturally adopted by a front seat passenger (if present) in a driving context.

Whilst it is clearly impractical to enforce drivers to recruit co-drivers to interact with them during long, monotonous or night-time journeys, advances in speech-recognition technologies, speech synthesis, natural language understanding (NLU) and dialogue management over recent years mean that modern voice-user interfaces (VUIs) are increasingly conversational – and often embodied by a digital personality: such technology may therefore provide a promising and viable alternative.

Moreover, given the success and popularity of digital assistants (e.g. Siri, Cortana and Alexa) in current personal devices such as smartphones, it is possible to envisage a future scenario where cars are embodied by similar digital assistants that interact with drivers using free-flowing, conversational dialogue akin to conversing with a loquacious passenger. Whilst such technology may be primarily intended to support the driving task by providing navigation advice, travel updates, vehicle status warnings, infotainment services etc., the inherent 'social' engagement afforded by interacting with the technology using speech [28] may also act to keep drivers alert and maintain driving performance.

1.4. Overview of Study

The aim of the study was to explore the effectiveness of a digital driving assistant as a countermeasure for driver fatigue. It was therefore important to create a situation in which drivers would begin to exhibit the symptoms of fatigue. During similar research (e.g. [10, 11]), a combination of time-on-task, time-of-day and driving monotony have been shown to induce fatigue-related driving symptoms, with effects revealed in a driving simulator during relatively short (40 minutes) drives [11] and afternoon testing [10], when circadian and homeostatic influences naturally reduce participants' alertness [29]. Adopting a similar approach to these studies, all testing was conducted in the afternoon between 14:00 and 16:00, and drivers were required to drive for 1 hour – 30 minutes accompanied by a digital driving assistant and 30 minutes alone (the order of conditions was counterbalanced to avoid learning effects). Based on the experience of earlier studies, a range of measures was employed to detect the onset and progression of driver fatigue (see Section 1.2).

2. Method

2.1. Participants

Twenty-three people took part in the study: 18 male and 5 female, with the median age range of 31-40 years old. All participants held a valid driving licence and were experienced and active drivers (mean time with licence: 11 years 10 months; current annual mileage, 7046). Participants were self-selecting volunteers who responded to advertisements placed on-line and around the University of Nottingham campus and were reimbursed with £20 (GBP) of shopping vouchers as compensation for their time. All participants provided written informed consent before taking part.

2.2. Apparatus, Design and Procedure

The study took place in a medium-fidelity, fixed-based driving simulator at the University of Nottingham (see Figure 1). The simulator comprises an Audi TT car located within a curved screen, affording 270 degrees forward and side image of the driving scene via three overhead HD projectors, together with rear and side mirror displays. A Thrustmaster 500RS force feedback steering wheel and pedal set are integrated with the original Audi steering wheel and pedals. The dashboard is created using a bespoke Java application and presented on a 7-inch Lilliput 668GL LCD screen, which replaces the original instrument cluster. The driving scenario was created using STISIM Drive (version 3) to replicate a standard three-lane UK motorway.

Participants undertook two drives, each lasting approximately 30 minutes. During each drive, participants were instructed to follow a lead vehicle, ‘as if travelling to a shared destination’, at a distance that they deemed to be safe and appropriate. The lead vehicle was present in front of the participant’s car at the start of the scenario and began moving in unison with the test participant’s vehicle. It remained in lane one of the motorway for the duration of both drives, travelling at a constant speed of 68-mph. During one of the drives, participants were provided with a digital driving assistant, described as “a fully operational, prototype natural language system currently under development by a major car manufacturer”. For the purpose of conducting the study, the digital assistant was created using a Wizard-of-Oz approach, in which a professional actor (Pablo) played the role of the assistant.

Fig. 1. *Medium-fidelity driving simulator showing motorway scenario used during study.*

The ‘Wizard of Oz’ system comprised a tablet computer, located within the driving simulator and connected to loudspeakers positioned inside the vehicle. Google Hangouts software was used to provide two-way voice and one-way video chat functionality between the participant and Pablo, who was situated out of sight in a nearby laboratory (see Figure 2) – Pablo could see and hear participants, whereas participants could only hear Pablo. To encourage participants to respond as if they were interacting with an autonomous system located within the car, rather than talking to another human at the end of a VOIP connection, the Google Hangouts interface was obscured, but the device itself remained in plain sight to provide a tangible entity and source for utterances: this misdirection was enhanced by overtly ‘installing’ the system in direct view of participants prior to them driving with the digital driving assistant (DDA). The DDA was called ‘David’, which was abbreviated to ‘Vid’ for the study, and introduced itself to participants before they began driving. Vid initiated conversation using the same opening gambits for each participant during the drive (see examples in Table 1), with all interjections requiring a response from participants; participants were told that they should interact with the system as they might should such a system exist in the real world. The intercourse therefore developed differently based on participant’s responses. Participants were also specifically told that they could initiate conversation with the DDA. System responses were generated in real time by Pablo, but were guided by a comprehensive script that detailed appropriate language and phrasing, and were delivered in a controlled fashion using a subtle computer inflexion, honed through extensive training and practice sessions, to mimic the prosodic contours associated with current state-of-the-art natural language speech interfaces. A second person accompanied Pablo and was responsible for retrieving facts and figures in real-time from the Internet and other sources, as dictated by the course of the conversation (e.g. locating requested music tracks, providing further details about news items etc.), and supplying these details to Pablo.

Table 1 Vid’s opening gambits

“I looked at your calendar and you have a meeting upcoming today at 3 o'clock. Would you like me to set a reminder for your meeting?”
“It looks like you've got a few things to do on your way home this evening. You need to buy milk. Would you like me to set a reminder for you to buy milk?”
“There is congestion ahead. This may delay you by 5 minutes would you like me to direct you around the congestion? ... OK. I am calculating a reroute.”

2.3. Measures

A range of objective and subjective performance measures (based on the aforementioned literature) were captured to assess driver alertness, workload, emotional state and vigilance, and temporal deterioration of driving performance; measures were subsequently compared across the two drives (identified as ‘control’ and ‘Vid’). The following measures are reported:

- Driving performance measures (lane keeping, speed and headway variability, response to hazard), captured from the driving simulator.

- Physiological measures (heart-rate variability and inter-beat interval using Empatica E4 wristband) captured throughout both drives
- Visual behaviour – blink duration and frequency, number and duration of fixations, spread of visual attention – percent road centre (PRC), pupil diameter
- Sleepiness ratings (Stanford Sleepiness Scale: [15]) captured before and after each drive
- Driver mood assessment (UWIST Mood Adjective Check-List: [30]) captured before and after each drive
- Workload ratings (NASA Task Load Index (TLX): [31]) captured before and after each drive

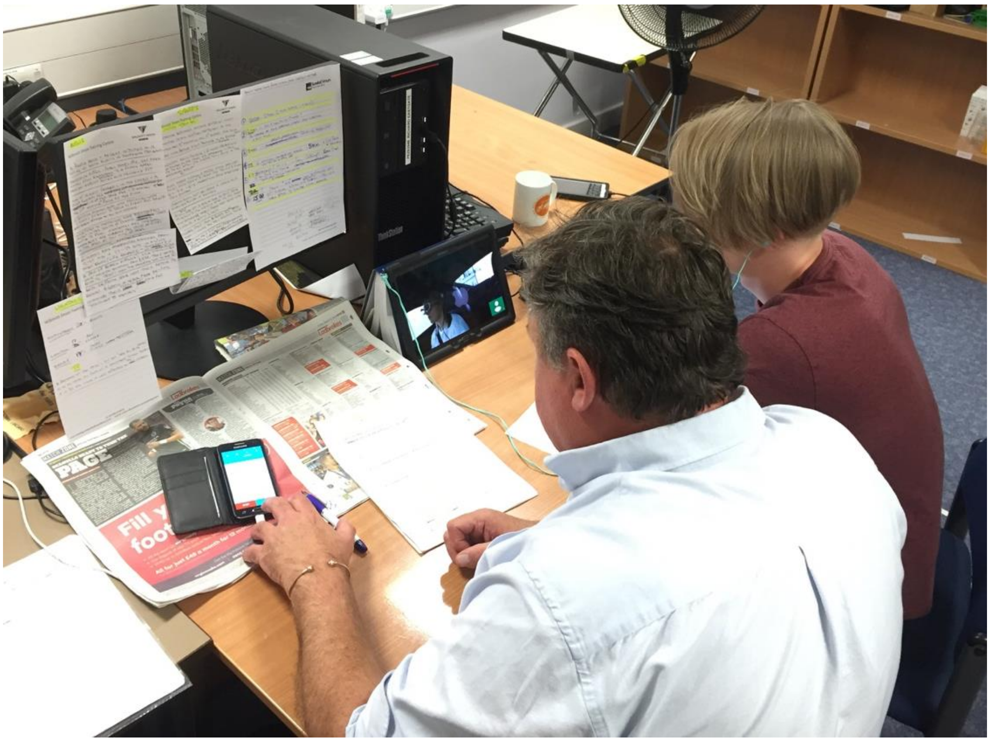

Fig. 2. *'Wizard of Oz' implementation of natural language system with actor ('wizard') (L), assistant (R) and tablet computer showing participant in driving simulator*.

3. Results and Analysis

3.1. Driving Performance

3.2.1 Lateral Control (Standard Deviation of Lane Positioning, SDLP): There was a significant difference in SDLP (t(18)=3.24, *p* = .005), with drivers demonstrating significantly more variability in lateral lane position during the control drive (M=0.93, SD=.28) compared to Vid (M=.814, SD=.259) (Figure 3).

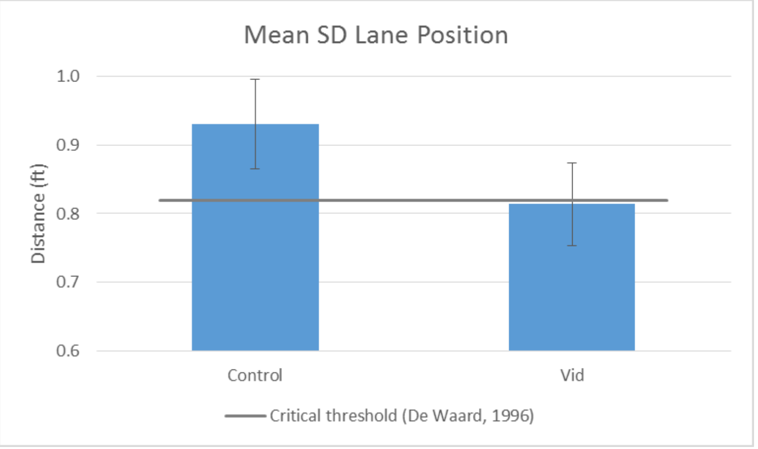

Fig. 3. *Mean standard deviation of lateral lane position, showing critical threshold identified by Brookhuis et al. [32]*

3.2.2 Longitudinal Control (Vehicle Speed, Headway): There were no significant differences in longitudinal vehicle control, with standard deviation of speed similar between conditions (p = .881). Differences in the standard deviation of headway were approaching significance (p = .101) with longer headway evident when driving with Vid (M=267.8, SD=563.0) compared to the control drive (M=226.6, SD=497.0).

3.2.3 Response to Hazard: A stationary vehicle was situated on the hard-shoulder of the motorway approximately 22.5 minutes into each drive. As participants approached the vehicle during their second drive, a pedestrian walked from behind the vehicle in the direction of the roadway, as if attending to or entering their vehicle. The headway to the vehicle at which drivers responded to this potential hazard (as indicated by them lifting their foot from the accelerator pedal) was recorded. On average, drivers responded 559 feet before the hazard following engagement with Vid (a simulated 'lost signal', prior to the hazard event ensured that drivers were not actually engaged with Vid at the time of the hazard, thereby avoiding any confounding distraction effects). In contrast, participants had already passed the pedestrian and parked car before lifting their foot from the accelerator pedal during the control drive (Figure 5).

3.1. Physiological Measure

3.2.1 Heart-Rate Variability: There was no significant difference in IBI or HRV between control (M=0.837, SD=.101 and M=0.068, SD=.020, respectively) and Vid (M=0.843, SD=.107 and M=0.067, SD=.021, respectively) drives (p = .518 and .631, respectively).

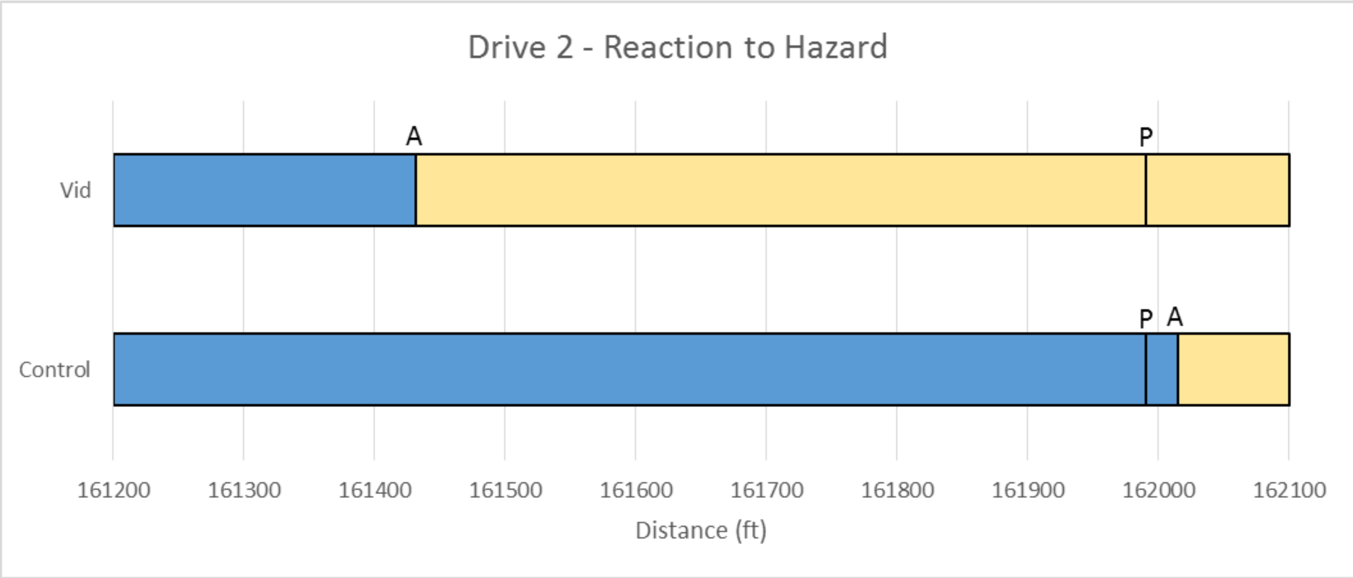

Fig. 4. *Mean distance at which drivers lifted their foot from the accelerator in response to a potential hazard (A = Participant's foot lifts from accelerator, P = Pedestrian begins walking)*

3.2. Eye Activity

3.3.1 Blinks: There was no significant difference in number of blinks between control (M=1097.9, SD=556.4) and Vid (M=1107.3, SD=424.6) drives (p = .918). There was also no significant difference in mean duration of blinks between control (M=397.4, SD=61.5) and Vid (M=383.5, SD=42.4) drives (p = .124).

3.3.2 Fixations: There was no significant difference in number of fixations between control (M=4020.8, SD=694.8) and Vid (M=4175.0, SD=839.6) drives (p = .434) and no significant difference in mean duration of fixations between control (M=255.9, SD=71.5) and Vid (M=249.7, SD=69.0) drives (p = .656).

3.3.2 Percentage Road Centre (PRC): There was a near significant difference in PRC between control (M=75.5, SD=12.2) and Vid (M=69.6, SD=21.6) drives (t(21)=1.91, p = .07), with participants directing more visual attention to the 'road centre' during the control drive (Figure 5).

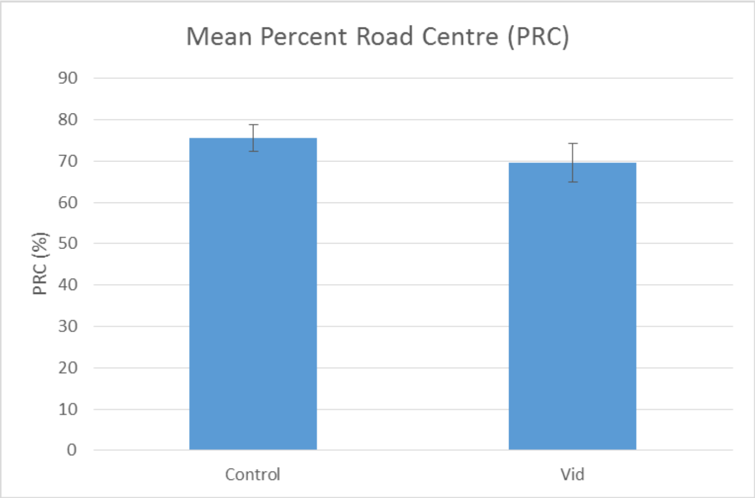

Fig. 5. *Mean percent road centre*

3.3.3 Visual Intake Mean Pupil Diameter: There was a significant difference in pupil size between control (M=2.79, SD=.89) and Vid (M=2.98, SD=.96) drives (t(25)=3.046, p = .005), with participants' pupil size greater during the 'Vid' drive (Figure 6).

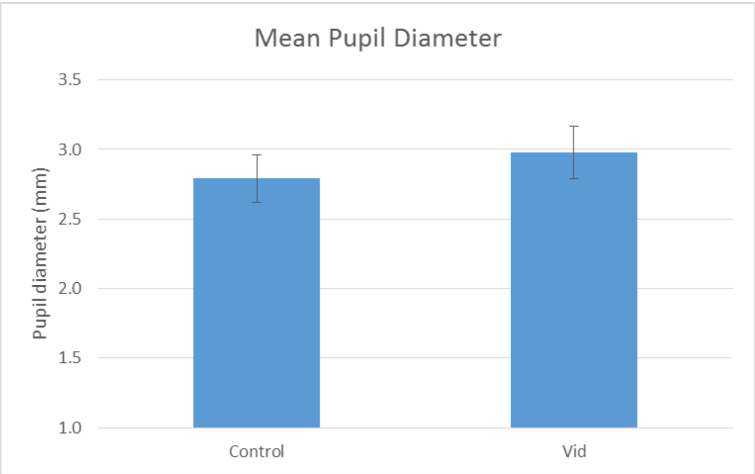

Fig. 6. *Mean pupil diameter*

3.3. Subjective Assessment

3.4.1 Sleepiness Ratings: There was a significant difference in sleepiness ratings following the control and Vid drives (t(21)=5.161, p < .005), with participants indicating more perceived sleepiness after the control drive (M=3.909, SD=1.477) compared to Vid (M=2.864, SD=1.457) (Figure 7).

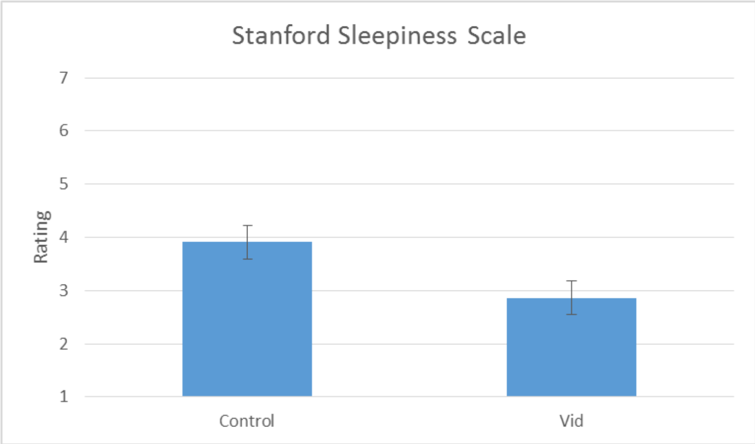

Fig. 7. *Mean sleepiness ratings*

1= Feeling active, vital, alert, or wide awake

2= Functioning at high levels, but not fully alert

3= Awake, but relaxed; responsive but not fully alert

4= Somewhat foggy, let down

5= Foggy; losing interest in remaining awake; slowed down

6= Sleepy, woozy, fighting sleep; prefer to lie down

7= No longer fighting sleep, sleep onset soon; having dream-like thoughts

3.4.2 Driver Mood Assessment: There was a significant difference in 'hedonic tone' and 'energetic arousal' following the control (M=6.76, SD=3.97 and M=-1.857, SD=4.15, respectively) and Vid (M=5.29, SD=4.23 and M=1.095, SD=3.62, respectively) drives (t(20)=2.353, p = .029 and t(20)=3.623, p = .002, respectively). Drivers indicated higher levels of 'alertness' (energetic arousal) and lower levels of 'happiness' (hedonic tone) associated with Vid. There was no significant difference in mood (tense arousal) between driving conditions (p = .557) (Figure 8).

3.4.3 Driver Workload: A paired samples t-test revealed that there was a significant difference in Total Workload between drives (t(22)=2.645, p = .015), with participants indicating lower perceived workload associated with the digital driving assistant (M=60.22, SD=36.4) compared to the control drive (M = 76.39, SD=37.3). Significant differences were evident in Mental (t(22)=2.855, p = .009) and Effort (t(22)=1.885, p = .073) subscales.

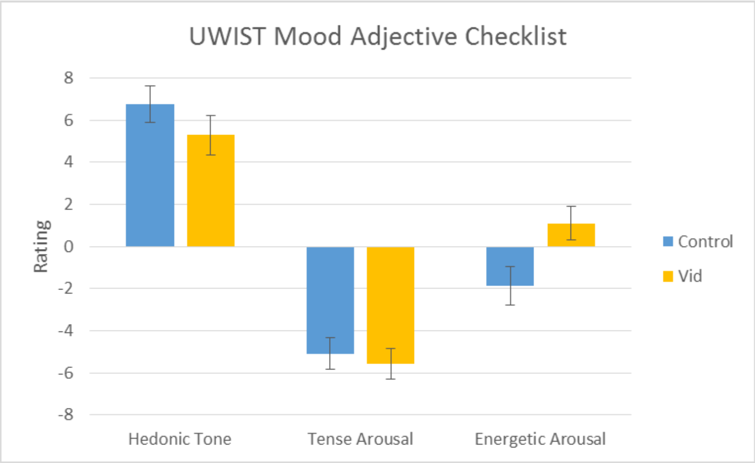

Fig. 8. *Mean UWIST mood adjective assessment ratings*

4. Discussion

The study investigated the effectiveness of interacting with an autonomous digital driving assistant as a countermeasure for driver fatigue. There is good evidence from the broad range of measures collected that the social interaction enabled by a natural language digital driving assistant ('Vid') had a positive effect on driver fatigue and arousal – minimising the effects of fatigue and maintaining arousal – compared to a similar drive during which no digital assistant was provided (the 'control' drive). In particular, driving with 'Vid' led to higher self-reported levels of alertness (energetic arousal), revealed by the UWIST Mood Adjective Checklist [30], and lower levels of subjective sleepiness, revealed by the Stanford Sleepiness Scale [15]. Objectively, driving performance was better throughout the drive in which participants interacted with the digital assistant, evidenced by better lane-keeping – significantly this was below the critical threshold identified by Brookhuis et al. [32], whereas SDLP for the control driver was higher – and earlier response to a potential hazard situation. Furthermore, there were physical indicators, such as increased pupil diameter associated with the digital driving assistant, suggesting that drivers were more alert when they were able to engage in conversation with the digital assistant; there were also lower levels of perceived workload associated with the 'Vid' drive, which we interpret as drivers' perceived effort to maintain alertness.

Nevertheless, these data should be treated with some caution – many objective measures revealed no differences between conditions, including physiological indicators (heart-rate variability and inter-beat interval). Eye activity measures, such as number and mean duration of blinks and fixations, and percent road centre, were equally inconclusive, although differences in the latter were nearing significance (p=.07). Furthermore, subjective measures, although apparently conclusive in the current study, are notoriously clumsy with respect to self-reported sleepiness. Evidence suggests that participants often overestimate their level of alertness with self-report, subjective measures, thereby concealing their fatigue [33]. Moreover, this discrepancy can be significant – a person might report low level of sleepiness (i.e. high level of alertness), yet carry an accumulated sleep debt with a high level of physiological sleepiness. In an environment stripped of factors that conceal the underlying physiological sleepiness (e.g. without 'Vid' to accompany them), such a person would be susceptible to the performance decrements associated with sleep loss [23].

It is also worth noting that given the experimental design (i.e. low feature driving scenario, monotonous driving task etc.), the results and discussions as presented relate only to *passive* task-related fatigue. Therefore, while the evidence suggests that conversational interactions can increase the attentional capacity of drivers in this situation (i.e. when experiencing cognitive *underload*), it is likely that similar interactions could add extra workload to drivers already highly-loaded (i.e. *actively* fatigued) – for example when negotiating high traffic density, complex road networks, poor weather etc. – and impair performance as a result. Consequently, such technology, *if intended to be used in the manner presented here*, must also consider driver workload prior to endorsing conversation as a viable countermeasure (as indeed a human passenger is likely to do).

There are inevitable difficulties in conducting fatigue-related driving studies in driving simulators, with some researchers attesting that drivers need to be driving for very lengthy periods of time, in addition to being sleep-deprived, to exhibit genuine fatigue. Nevertheless, this is not necessarily supported by real-world evidence [18]. Moreover, our intention was to provide a context in which low arousal (symptomatic of fatigue) was likely to arise among participants and we therefore conducted testing in the afternoon and utilised a low-feature, monotonous driving scenario. Such conditions have been shown to induce fatigue-related driving symptoms in driving simulators [11, 10, 29]. Nevertheless, the lack of conclusive findings in some key measures (particularly eye activity) may suggest that such conditions were actually insufficient to induce genuine fatigue in all participants. It is also possible that the increased arousal during the 'Vid' drive – revealed in other measures – may have been due to the novelty of the situation. A further potential limitation may be revealed in the gender-skewed sample (male:female, approximately 3:1). Although this is a natural symptom of convenience sampling, responses to the digital driving assistant (and therefore its implied success as a countermeasure to driver fatigue) may have varied between genders, particularly given previous research suggesting that the 'gender' of voice user interface agents can influence attitudes based on an interlocutor's gender [28].

As a final cautionary note, it is worth highlighting that operational countermeasures (such as social interaction with a digital driving assistant), do not necessarily address the underlying physiological causes of fatigue, but rather aim to enhance alertness and performance temporarily so that operational safety and performance are maintained [23]. Moreover, individual differences – such as age, daytime sleep propensity, and lifestyle factors such as new parenthood, personality and mood – can also play a major part in susceptibility to fatigue. Consequently, driver fatigue is a difficult risk to manage. While the current evidence suggests that social interaction with a digital driving assistant may act as an effective countermeasure to passive task-related fatigue in some situations, it is important for drivers to understand the need to individually assess their own fatigue, both before and during driving.

5. Conclusions

A driving simulator study has revealed that social interaction and conversation with an autonomous digital driving assistant may act as an effective countermeasure to driver fatigue. Nevertheless, this conclusion is based on the device as presented during the study – current 'voice' technology is unlikely to offer such capability for a number of years. In addition, several key measures were inconclusive. Further investigations should therefore accompany the development of 'voice' technology, and consider longer episodes of driving with possible sleep-deprivation of participants prior to testing. Additionally, the effect of social interaction on drivers already highly-loaded (i.e. *actively* fatigued) could be considered.

6. Acknowledgements

The research was conducted in collaboration with Jaguar Land Rover Research and the authors gratefully acknowledge their support. We would also like to thank Vicki Antrobus, Ayse Eren, Harry Large and our wizard, Pablo Raybould, without whom the research would not have been possible.

7. References

[1] P. Jackson, C. Hilditch, A. Holmes, N. Reed, N. Merat and L. Smith, "Fatigue and road safety: a critical analysis of recent evidence," London: Department for Transport, 2011.

[2] D. Dinges, "An overview of sleepiness and accidents," *Journal of Sleep Research,* vol. 4, no. 2, pp. 4-14, 1995.

[3] E. Grandjean, "Fatigue in industry.," *British Journal of Industrial Medicine,* vol. 36, pp. 175-186, 1979.

[4] L. Torsvall and T. Åkerstedt, "Sleepiness on the job: continuously measured EEG changes in train drivers," *Electroencephalography and clinical Neurophysiology,* vol. 66, no. 6, pp. 502-511, 1987.

[5] J. A. Horne and L. A. Reyner, "Sleep related vehicle accidents," *British Medical Journal,* vol. 310, pp. 565-567, 1995.

[6] S. K. Lal and A. Craig, "A critical review of the psychophysiology of driver fatigue," *Biological psychology,* vol. 55, no. 3, pp. 173-194, 2001.

[7] P. Philip, P. Sagaspe, N. Moore, J. Taillard, A. Charles, C. Guilleminault and B. Bioulac, “Fatigue, sleep restriction and driving performance,” *Accident Analysis and Prevention,* vol. 37, no. 3, pp. 473-478, 2005.

[8] S. Adkins, “Performance, heart rate, and respiration rate on the day-night continuum,” *Perceptual and motor skills,* vol. 18, no. 2, pp. 409-412, 1964.

[9] K. Brookhuis and D. de Waard, “Monitoring drivers’ mental workload in driving simulators using physiological measures,” *Accident Analysis & Prevention,* vol. 42, no. 3, pp. 898-903, 2010.

[10] S. Otmani, T. Pebayle, J. Roge and A. Muzet, “Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers,” *Physiology and Behavior,* vol. 84, no. 5, pp. 715-724, 2005.

[11] P. Thiffault and J. Bergeron, “Monotony of road environment and driver fatigue: a simulator study,” *Accident Analysis and Prevention,* vol. 35, no. 3, pp. 381-391, 2003.

[12] W. Wierwille and L. Ellsworth, “Evaluation of driver drowsiness by trained raters,” *Accident Analysis & Prevention,* vol. 26, no. 5, pp. 571-581, 1994.

[13] T. W. Victor, J. L. Harbluk and J. A. Engström, “Sensitivity of eye-movement measures to in-vehicle task difficulty,” *Transportation Research Part F: Traffic Psychology and Behaviour,* vol. 8, no. 2, pp. 167-190, 2005.

[14] M. W. Johns, “A new method for measuring daytime sleepiness: the Epworth sleepiness scale,” *Sleep,* vol. 14, no. 6, pp. 540-545, 1991.

[15] E. Hoddes, V. Zarcone, H. Smythe, R. Phillips and W. C. Dement, “Quantification of sleepiness: a new approach,” *Psychophysiology,* vol. 10, no. 4, pp. 431-436, 1973.

[16] S. N. Biggs, A. Smith, J. Dorrian, K. Reid, D. Dawson, v. d. H. C. and S. Baulk, “Perception of simulated driving performance after sleep restriction and caffeine,” *Journal of Psychosomatic Research,* vol. 63, no. 6, pp. 573-577, 2007.

[17] S. D. Baulk, S. N. Biggs, K. J. Reid, C. J. van den Heuvel and D. Dawson, “Chasing the silver bullet: measuring driver fatigue using simple and complex tasks,” *Accident Analysis and Prevention,* vol. 40, no. 1, pp. 396-402, 2008.

[18] N. B. Powell, K. B. Schechtman, R. W. Riley, K. Li, R. Troell and C. Guilleminault, “The road to danger: the comparative risks of driving while sleepy,” *Laryngoscope,* vol. 111, no. 5, pp. 887-893, 2001.

[19] G. Larue, A. Rakotonirainy and A. Pettitt, “Driving performance impairments due to hypovigilance on monotonous roads,” *Accident Analysis & Prevention,* vol. 43, no. 6, pp. 2037-2046, 2011.

[20] J. Santamaria and K. Chiappa, “The EEG of drowsiness in normal adults,” *Journal of Clinical Neurophysiology,* vol. 4, no. 4, pp. 327-382, 1987.

[21] J. Horne and L. Reyner, “Counteracting driver sleepiness: effects of napping, caffeine, and placebo,” *Psychophysiology,* vol. 33, no. 3, pp. 306-309, 1996.

[22] Department for Transport, “Wake Up to the Signs of Tiredness. THINK! leaflet,” 2005. [Online]. Available: http://webarchive.nationalarchives.gov.uk/20061101004827/http://thinkroadsafety.gov.uk/advice/pdf/992.pdf. [Accessed 22 August 2016].

[23] M. R. Rosekind, P. H. Gander, K. B. Gregory, R. M. Smith, D. L. Miller, R. Oyung, L. L. Webbon and J. M. Johnson, “Managing fatigue in operational settings 1: Physiological considerations and countermeasures,” *Behavioral Medicine,* vol. 21, no. 4, pp. 157-165, 1996.

[24] N. Wright, B. Stone, T. Horberry and N. Reed, “A review of in-vehicle sleepiness detection devices,” 2007.

[25] J. A. Caldwell, M. M. Mallis, J. L. Caldwell, M. A. Paul, J. C. Miller and D. F. Neri, “Fatigue countermeasures in aviation,” *Aviation, space, and environmental medicine,* vol. 80, no. 1, pp. 29-59, 2009.

[26] M. Dijkman, N. Sachs, E. Levine, M. Mallis, M. Carlin, K. Gillen, J. Powell, S. Samuel, J. Mullington, M. Rosekind and D. Dinges, “Effects of reduced stimulation on neurobehavioral alertness depend on circadian phase during human sleep deprivation,” *Sleep Res,* vol. 26, p. 265, 1997.

[27] P. Cabon, J. Fouillot, A. Coblentz, R. Mollard and J. Speyer, “Etude du niveau d'éveil des équipages au cours de vols long-courriers,” *Medicine Aéronautique et Spatiale,* vol. 30, pp. 321-338, 1991.

[28] C. Nass and S. Brave, Wired for speech: How voice activates and advances the human-computer relationship, Cambridge: MIT press, 2005.

[29] P. H. Ting, J. R. Hwang, J. L. Doong and M. C. Jeng, “Driver fatigue and highway driving: A simulator study,” *Physiology & Behavior,* vol. 94, no. 3, pp. 448-453, 2008.

[30] G. Matthews, D. Jones and A. Chamberlain, “Refining the measurement of mood: The UWIST mood adjective checklist,” *British journal of psychology,* vol. 81, no. 1, pp. 17-42, 1990.

[31] S. Hart and L. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” *Advances in psychology,* vol. 52, pp. 139-183, 1988.

[32] K. Brookhuis, D. Waard and S. Fairclough, “Criteria for driver impairment,” *Ergonomics,* vol. 46, no. 5, pp. 433-445, 2003.

[33] D. F. Dinges, “The nature of sleepiness: causes, contexts and consequences,” *Perspectives in behavioral medicine: Eating, sleeping, and sex,* pp. 147-179, 1989.