Assessment of police subjective workload and preference for using a voice-based interface during simulated driving

Mitsopoulos-Rubens, E., Filtness, A. & Lenné, M. G.

Monash University Accident Research Centre (MUARC), Monash University, Australia

Abstract

Police in-vehicle systems include a visual output mobile data terminal (MDT) with manual input via touch screen and keyboard. This study investigated the potential for voice-based input and output modalities for reducing subjective workload of police officers while driving. Nineteen experienced drivers of police vehicles (one female) from New South Wales (NSW) Police completed four simulated urban drives. Three drives included a concurrent secondary task: an imitation licence number search using an emulated MDT. Three different interface output-input modalities were examined: Visual-Manual, Visual-Voice, and Audio-Voice. Following each drive, participants rated their subjective workload using the NASA - Raw Task Load Index and completed questions on acceptability. A questionnaire on interface preferences was completed by participants at the end of their session. Engaging in secondary tasks while driving significantly increased subjective workload. The Visual-Manual interface resulted in higher time demand than either of the voice-based interfaces and greater physical demand than the Audio-Voice interface. The Visual-Voice and Audio-Voice interfaces were rated easier to use and more useful than the Visual-Manual interface, although not significantly different from each other. Findings largely echoed those deriving from the analysis of the objective driving performance data. It is acknowledged that under standard procedures, officers should not drive while performing tasks concurrently with certain in-vehicle policing systems; however, in practice this sometimes occurs. Taking action now to develop voice-based technology for police in-vehicle systems has potential to realise visions for potentially safer and more efficient vehicle-based police work.

Introduction

For many police officers, the police vehicle operates as their workstation/office for much of their work shift. In a survey of 476 officers from NSW Police, approximately 75% indicated spending at least half of their shift in a police vehicle, with about 70% of these officers reporting that at least three quarters of their shift is spent in a police vehicle (Mitsopoulos-Rubens, Trotter & Rudin-Brown, 2009). Through the mobile data terminal (MDT), officers can access several systems and databases which are used for police work. However, in many respects, the MDT looks and operates like an office computer. The typical MDT comprises a visual display with touch screen capabilities. There is also a separate keyboard connected to a unit in the glove box. For some tasks, the keyboard is generally preferred as issues with touch screen sensitivity and the small size of icons may lead to inefficiencies in inputting the required information to the MDT were the touch screen to be used (Mitsopoulos-Rubens, Filtness & Lenné, 2013).

Thus, in general, the design of the MDT interface has not been optimised for the in-vehicle work environment of a police officer. This has potential negative implications, not only for policing operations, but for driver and passenger safety. While officers are advised not to interact with the MDT while driving, in practice, interaction with the MDT while driving may occur. This is particularly likely in situations where there is no front seat passenger – that is, in situations where the driver is the only police officer in the vehicle (Hampton & Langham, 2005).

For the general driving population, there is much evidence on the distracting potential of in-vehicle information and communication systems (e.g., Klauer, Dingus, Neale, Sudweeks & Ramsey, 2006; Mitsopoulos-Rubens, Trotter & Lenné, 2011; Young, Mitsopoulos-Rubens, Rudin-Brown & Lenné, 2012). This potential is at its highest for visually demanding interfaces involving direct manipulation – that is, interfaces that require drivers to take their eyes off the road and hands off the steering wheel. For this reason, there is much interest in the use of voice-based interfaces as a mechanism through which to alleviate at least some of the negative effects on driving performance and subjective workload associated with visual-manual interfaces (e.g., Garay-Vega, et al., 2010; Lee, Caven, Haake & Brown, 2001; Maciej & Vollrath, 2009; Ranney, Harbluk & Noy, 2005).

Police officers receive specialised driver training to enable them to operate a vehicle under conditions of higher workload (e.g., high speeds) and competing task demands. However, it is not known whether police officers are more adept than individuals in the general driving population at managing the added demands associated with concurrent operation of an in-vehicle device while driving. This is a critical gap in knowledge given the current interaction requirements of the MDT and that drivers of police vehicles may, depending on the prevailing circumstances, need to interact with the MDT while driving. Moreover, whether a voice-based interface would confer any advantages in use over a traditional visual-manual interface given the sorts of in-vehicle tasks which police officers might typically undertake (e.g., person and licence number checks/searches) with the MDT has yet to be established. Thus, in the absence of any direct evidence, it cannot be assumed that experienced drivers of police vehicles will respond to interaction with an in-vehicle device – whether it be visual-manual or voice-based – concurrently while driving in exactly the same way as experienced drivers in the general driving population. In other words, there is a need to establish whether the potential for driver distraction is greatest among police officers interacting with a visual-manual interface while driving than police officers using a voice-based interface while driving. Were a voice-based interface found to be less distracting, then it suggests that implementing a voice-based interface in police vehicles may be an appropriate option to pursue, potentially enabling safer and more efficient vehicle-based police work.

A simulator study was undertaken to examine the effects on operations, including driving, of using a voice-based interface for in-vehicle policing tasks compared with the traditional MDT visual-manual interface. Participants were experienced drivers of police vehicles. Two voice-based interface variants were examined in order to determine whether there are any advantages associated with a complete voice-based system (i.e., both input and output) over a partial voice-based system (i.e., input or output, but in this case, input). In both cases, inputs to the system were through voice. In one case, the system provided information to the user through a visual display, while in the other system information was provided to the user in the form of audio output. The current paper reports on those aspects of the study concerned with the assessment of subjective workload, interface preferences and acceptability. Therefore, of particular interest here is whether the voice-based interface types are associated with lower levels of subjective workload, higher levels of acceptability, and preferred by more experienced drivers of police vehicles than the traditional MDT visual-manual interface. Whether a given voice-based interface confers any advantages over the other in terms of subjective workload, acceptability and/or preferences was also of interest.

Method

Prior to undertaking this research, ethics approval was obtained from the Monash University Human Research Ethics Committee.

Design

The study used a dual-task paradigm. The primary task was a driving task and the secondary task was a licence plate task. A single, within-subjects, independent variable characterized measures relating directly to the driving task (e.g., subjective workload). Labeled “condition”, this variable comprised four levels: Baseline, Visual-Manual, Visual-Voice and Audio-Voice. The latter three levels refer to the three interface designs under study.

Participants

Participants were 19 (18 males and 1 female) experienced drivers of police vehicles. Participants were in the age range of 35 to 57 years (Mean = 47.2 years; SD = 6.8). All participants held a full car driver's licence, and had been licensed to drive a car for at least 18 years (Mean = 30.1 years; SD = 6.9). All participants were current employees of NSW Police, and had worked for NSW Police for at least five years (Mean = 18.0 years; SD = 8.4 years). Participants reported spending an average of 20.7 hours (SD = 9.2) each week in a police vehicle.

Tasks and equipment

The MUARC Driver Distraction Test (DDT; Young, Lenné, Archer & Williamson, 2009) was used in the current study. The test is a driving simulation spanning approximately 6.6 kilometres of straight, undivided road set in an urban environment. In general, participants' task is to drive as they would normally, taking into consideration the speed limit and other road users. Participants are required to travel in the left lane, unless directed by signs to change lanes or to turn, and to not overtake other vehicles. Each run of the test takes approximately 10 minutes to complete. The test was implemented in the MUARC mid-range driving simulator, using Eca Faros driving simulation software. Briefly, the simulator consists of a full-size Holden sedan on a fixed-base. The vehicle is surrounded by two projection screens. The front, curved screen provides a field of view of approximately 180 degrees horizontally. The rear, flat screen provides a field of view of about 60 degrees horizontally. Projectors convey images of pre-programmed traffic scenes onto the screens, and a sound system provides the driver with realistic traffic sounds. The experimenter runs and monitors the simulations from the control room, which is located adjacent to the simulator room.

A licence number search task was developed for use in this study as the secondary task. This task was based on the comparable task performed by police officers quite routinely when in their vehicles. Thus, a licence number task was considered a suitable candidate for use in the current study. Three versions of the task were developed: one for each of the three interface types under study: Visual-Manual, Visual-Voice and Audio-Voice. In every case, the task was programmed and run using DirectRT v2012 experimental psychology software (Empirisoft Corporation). Item stimuli comprised a six-character licence plate number in the form of two letters, two numbers, and two letters (e.g., BM48RP). This configuration is consistent with that being implemented in NSW at the time of the study. In general, participants' task for each item was to repeat the licence plate number when prompted to do so, and to then wait for the outcome of the search (i.e. "Match" or "No match"). In responding to each stimulus, participants were advised that their goal should be to respond as quickly as they felt they needed to, but not so quickly that they made lots of errors. In every version of the task, items were presented in a random order from a list of 60 items. Further, item administration was self-paced – that is, the next item was not presented until the current item had been completed. Once an item had been completed, the stimulus for the next item was presented following a delay of 10 seconds. The stimulus presentation, output mode (i.e., visual or audio) and the response, input mode (i.e., manual or voice) varied according to the interface type. A brief description of each interface type, and how the task was implemented in every case, is given in Table 1.

While intending to mimic the use of speech recognition software, the Visual-Voice and Audio-Voice task versions were not implemented in this way. This was to ensure that the outcomes of the three task versions were directly comparable. The task versions were implemented by having a second experimenter manually enter participants' responses as the responses were given. The second experimenter was seated in the control room and out of participants' view.

Questionnaires

In addition to a short background (demographic, driving experience, work experience) questionnaire, participants completed three questionnaires as part of the study: a "subjective workload" questionnaire, an "acceptability" questionnaire, and a "preferences" questionnaire.

The NASA – Raw Task Load Index (NASA-RTLX) (Byers, Bittner and Hill, 1989) was used to provide an indication of participants' perceived level of workload associated with the driving task. The NASA-RTLX is a multi-dimensional rating scale that provides a numerical score (from "0" to "100") for each of six workload dimensions: mental demand, physical demand, time demand, performance, effort, and frustration level. Each dimension uses a visual analogue scale.

Table 1. Brief description of the three interface types explored in the study

Interface type	**Output**	**Input**	**Task item composition**
Visual-Manual	Visual through the display (see Figure 1)	Manual through the touch screen and separate keyboard	1. "Beep" sound to alert participants to look at the display. 2. Licence number presented on the display for 5 seconds. 3. Response box appeared on the display. Participants' task was to type the licence number and when done to press the on-screen "OK" button. 4. For 2 seconds, text "Searching…" appeared on the display. 5. For 1.5 seconds, text "Match" or "No match" shown on the display.
Visual-Voice	Visual through the display (see Figure 1)	Voice through microphone (attached to headset worn by participants)	1. "Beep" sound to alert participants to look at the display. 2. Licence number presented on the display for 5 seconds. 3. Response box appeared on the display. Participants' task was to say the licence number out loud and when done to say the word "enter". 4. For 2 seconds, text "Searching…" appeared on the display. 5. For 1.5 seconds, text "Match" or "No match" shown on the display.
Audio-Voice	Audio through the headset	Voice through microphone (attached to headset worn by participants)	1. "Beep" sound to alert participants to look at the display. 2. Pre-recorded licence number presented through the headset. Took 5 seconds to play. 3. Participants' task was to say the licence number out loud and when done to say the word "enter". 4. After 2 seconds, the pre-recorded term "Match" or "No match" was presented through the headset. Took 1.5 seconds to play.

User acceptability is an important determinant of system use. To enable direct comparison of the acceptability of the three interfaces, a variant of the questionnaire developed by Davis (1989) was used. This questionnaire has two scales, with each scale comprising six items. The two scales distinguish between two facets of acceptability: usefulness and ease of use. Davis (1989) defines perceived usefulness as the extent to which individuals believe that using a given system would improve their job performance. He defines perceived ease of use as the extent to which individuals believe that using a given system would be free of effort. Participants responded to each item by placing a line at the appropriate point on a visual analogue scale, with end-points "likely" and "unlikely". In this way, each item could be given a numerical score from 0 (i.e., "likely") to 100 (i.e., "unlikely"). For each scale, an overall score was derived for each participant by averaging responses across the six items making up the questions for that scale.

To provide an indication of participants' preferences regarding interface type, a questionnaire was developed that asked participants to rank the three interfaces on each of several dimensions. The dimensions were: efficiency, satisfaction, comfort, feasibility, and driving safety. Participants were also asked to rank the interfaces according to their preferences overall. In responding to each question, participants were asked to think about police operations and the range and sorts of communication and interaction tasks that are usually carried out in police vehicles as part of police work.

Figure 1. Location of the visual display (left) used for the licence number task

Procedure

Sessions were conducted on an individual basis, with each session lasting for approximately two hours. Following a suite of preparatory, training and practice exercises, the experimental trials were conducted. These comprised four sets of trials: a baseline driving trial, and three sub-sets of trials involving the licence number task - that is, one sub-set for each of the three interface types under study. For each trial sub-set, participants first completed a baseline trial of the licence number task, followed by a trial in which participants performed the licence number task in conjunction with the driving task. Participants were instructed to prioritise the driving task in every case. Following every driving trial, participants completed the subjective workload questionnaire, and, if relevant, the appropriate version of the acceptability questionnaire. To minimize order effects, presentation of the four sets of trials was counterbalanced across participants. Further, half of the participants experienced the trial sets using one order of the four available DDT drives and half using a different order. To conclude their session, participants completed the final questionnaire on interface preferences.

Results

In preparation for analysis, the subjective workload and acceptability data were screened for outliers and violations of normality. Neither outliers nor normality violations were identified. These data, in turn, were analysed using either analysis of variance (ANOVA) or t-tests. Statistical significance was defined as $p \leq 0.05$. A significant ANOVA result was followed with appropriate post-hoc tests. Bonferroni adjustment was applied to control for potential Type I errors due to multiple comparisons. The preferences data were analysed using chi-square. Again, statistical significance was defined as $p \leq 0.05$.

Subjective workload

Mean scores on each of the six dimensions of the NASA-RTLX are presented in Table 2. The score is given separately for each of the four driving conditions. A one-way repeated measures ANOVA conducted separately for each dimension revealed, in every case, a significant effect of condition (Table 2). To locate the source of each significant effect, a series of six paired sample t-tests with adjusted p-value ($p \leq 0.008$) was carried out for each dimension. For each of the six dimensions, mean ratings were significantly lower for the Baseline condition than for the Visual-Manual and Visual-Voice conditions. A slightly different pattern emerged for the Audio-Voice condition: with the exception of physical demand, mean scores in the Baseline condition were significantly lower than for the Audio-Voice condition for each dimension. In the case of physical demand, mean ratings did not differ significantly between the Baseline and Audio-Voice conditions.

Further significant differences between conditions were found in the case of physical demand and time demand. Participants reported perceiving a significantly higher level of physical demand and time demand when using, while driving, the Visual-Manual interface than the Audio-Voice interface. Time demand was also perceived to be higher for the Visual-Manual interface than for the Visual-Voice interface.

Table 2. Mean subjective workload scores for each of the four driving conditions and ANOVA results (Note. With the exception of Performance, lower score is better)

Dimension	Baseline	Visual-Manual	Visual-Voice	Audio-Voice	ANOVA result
Mental demand	28.47	79.37	62.58	63.21	$F(3,54)=21.26$*
Physical demand	20.84	63.00	41.74	26.11	$F(3,54)=17.13$*
Time demand	22.95	73.10	49.47	43.95	$F(3,54)=21.12$*
Performance	24.74	71.05	54.05	53.84	$F(3,54)=17.64$*
Effort	26.74	79.37	58.26	59.58	$F(3,54)=20.19$*
Frustration	24.74	63.16	44.63	46.16	$F(3,54)=8.99$*

*$p \leq 0.001$

Acceptability

In the current research, acceptability was defined as comprising two dimensions: ease of use and usefulness. For each dimension, an overall score was derived for each participant by averaging responses across the six relevant questionnaire items making up that dimension. Table 3 presents the mean scores obtained for each dimension as a function of interface type.

In the case of usefulness, participants' ratings for the Visual-Voice and Audio-Voice interfaces were made by relating usefulness (as operationalised in a given question) to that for the Visual-Manual interface. For this reason, no usefulness scores are available for the Visual-Manual interface (see Table 3). A paired sample t-test revealed that, relative to the Visual-Manual interface, participants did not perceive the Visual-Voice interface to be any more or less useful than the Audio-Voice interface type.

Table 3. Mean usefulness and ease of use scores for each interface type and statistical test results (Note. Lower score is better)

Dimension	Visual-Manual	Visual-Voice	Audio-Voice	Statistical test result
Usefulness	-	22.64	24.40	F(3,54)=21.26*
Ease of use	61.91	21.18	20.81	F(3,54)=17.13*

*$p \leq 0.001$

With regards to ease of use, a one-way repeated measures ANOVA revealed a significant effect of interface type (Table 3). A series of paired sample t-tests with adjusted p-value ($p \leq 0.017$) showed that participants perceived the Visual-Manual interface to be less easy to use overall than either the Visual-Voice interface or the Audio-Voice interface. Perceived ease of use was not found to differ significantly between the Visual-Voice and the Audio-Voice interface types.

Preferences

Table 4 summarises, for each dimension, the outcomes of the preference questionnaire. For every dimension, including overall, the Visual-Manual interface was overwhelmingly the least favoured by participants. The results were less clear about which interface, Visual-Voice or Audio-Voice, participants preferred the most. Overall, just over half of the participants expressed a preference for the Audio-Voice interface and just under half indicated that they preferred the Visual-Voice interface.

Table 4. Interface rankings and chi-square test result for each dimension (Note. Values are percentages; For a given dimension and rank, the highest percentage over 50% is emboldened)

Dimension	Interface type	Rank 1st (best)	Rank 2nd	Rank 3rd (worst)	Chi-square test result
Efficiency	**Visual-Manual**	0	10.5	**89.5**	$\chi^2(4)=43.90$*
	Visual-Voice	42.1	**57.9**	0	
	Audio-Voice	**57.9**	31.6	10.5	
Satisfaction	**Visual-Manual**	0	10.5	**89.5**	$\chi^2(4)=42.00$*
	Visual-Voice	47.4	**52.6**	0	
	Audio-Voice	**52.6**	36.8	10.5	
Comfort	**Visual-Manual**	0	10.5	**89.5**	$\chi^2(4)=43.90$*
	Visual-Voice	42.1	**57.9**	0	
	Audio-Voice	**57.9**	31.6	10.5	
Feasibility	**Visual-Manual**	0	10.5	**89.5**	$\chi^2(4)=41.37$*
	Visual-Voice	**52.6**	47.4	0	
	Audio-Voice	47.4	42.1	10.5	
Safety	**Visual-Manual**	0	5.3	**94.7**	$\chi^2(4)=53.37$*
	Visual-Voice	36.8	47.4	0	
	Audio-Voice	**63.2**	42.1	5.3	
OVERALL	**Visual-Manual**	0	10.5	**89.5**	$\chi^2(4)=42.00$*
	Visual-Voice	47.4	**52.6**	0	
	Audio-Voice	**52.6**	36.8	10.5	

*$p \leq 0.001$

Regarding specific dimensions, a slightly larger proportion of participants assigned the highest ranking to the Audio-Voice interface than the Visual-Voice interface in every case, particularly safety. That is, slightly more participants felt that the Audio-Voice interface would be the least likely to impact negatively their ability to drive a police vehicle safely. The only exception to this general pattern of findings was the feasibility dimension: a slightly higher proportion of participants – that is, just over half - indicated that, given the nature of in-vehicle policing tasks, the Visual-Voice interface would be the most feasible option.

Discussion and conclusions

The overarching objective of the research presented in this paper was to examine the effects on operations of interacting with a voice-based interface while driving relative to those effects associated with use of a visual-manual interface while driving. The focus in the current paper was on subjective workload, preferences and acceptability. The effects on driving performance and secondary (licence number) task performance are presented in detail elsewhere (see Mitsopoulos-Rubens et al., 2013).

Participants' acceptability scores and preferences showed a very consistent pattern. In every case, the Visual-Manual interface was deemed the least acceptable and was rated as the least preferred of the three interface types under study. In terms of acceptability, neither voice-based interface was rated as either more or less easy to use and useful than the other. In the case of the preferences questionnaire, participants were required to rank one interface over another. While participants showed a slight preference overall for the Audio-Voice interface over the Visual-Voice interface, participants, in general, appeared to be divided as to which voice-based interface they preferred the most. Regarding the specific dimensions of interest, more than half of the participants (up to 63%) assigned a higher ranking to the Audio-Voice interface than to the Visual-Voice interface. That is, slightly more participants felt that the Audio-Voice interface than the Visual-Voice interface would allow them to perform police operations involving the MDT more efficiently, and under conditions of greater comfort, satisfaction, and, critically, safety. These results in favour of the Audio-Voice interface are most likely a reflection of participants' desire, at least in principle, to divert their visual attention away from the roadway as little as is possible. It is interesting that a slightly larger proportion of participants ranked as their most preferred interface the Visual-Voice interface in terms of feasibility. That is, despite a slight, general preference for the interface which demands the least amount of eyes off road time, a slightly larger proportion of participants felt that the Visual-Voice interface would be the more practical solution given the sorts and range of information and communication tasks undertaken by police officers using the MDT in the in-vehicle context.

In responding to the acceptability and preferences questionnaires, participants were asked to draw on their experiences with the MDT and knowledge of in-vehicle police operations. Given that the visual-manual interface is the type of interface in current use, participants could draw on a larger and more detailed database of experience in providing their responses for the Visual-Manual interface than for either the Audio-Voice or Visual-Voice interface types. It is not known to what extent these different levels of experience across interface types influenced acceptability scores and preferences. As they stand, the acceptability and preferences results show a clear preference for voice-based interfaces, in general, over visual-manual interface types.

Arguably, subjective workload scores were less influenced by participants' prior experiences with a given interface type as, in responding to the items of the NASA – RTLX, participants were asked to reflect on the drive just performed. For each dimension of the NASA – RTLX, participants reported perceiving a significantly higher level of workload in each of the dual-task conditions (i.e., performing the licence number task while driving) relative to the Baseline (i.e., just driving). There was one exception, however: there was no significant difference between the Baseline and Audio-Voice conditions in the perceived level of physical demand. Thus, for the most part, performing a secondary task while driving resulted in increased levels of subjective workload. Of further interest here was whether there were any significant differences in subjective workload between interface types. Critically, some differences of this sort were observed. Between interface types, the level of physical demand experienced in the Visual-Manual condition was reported as being significantly higher than that experienced in the Audio-Voice condition. Further, participants reported perceiving a greater degree of time pressure while driving in the Visual-Manual condition than in both the Audio-Voice and Visual-Voice conditions. These data suggest that there are some potential gains in subjective workload to come from the use of voice-based interfaces as opposed to interfaces of the visual-manual type. However, that there was no difference between interface types in terms of the perceived level of mental demand is an interesting one as it highlights that the use of voice-based interfaces is still associated with imposing significant cognitive load. These general finding are consistent with those of others. Ranney et al. (2005), for example, did not find a difference in subjective workload ratings between a visual-manual system and voice-based (i.e., visual-voice) system. Lee et al. (2001) found significantly higher subjective workload estimates under driving conditions when a speech-based email system was used than when it was not.

Moreover, Lee et al. (2001) reported that subjective workload was most affected when the task being performed with the speech-based email system was a “complex” one than when it was a “simple” one. Participants also rated the complex scenario as more distracting than the simple scenario. In using the complex version of the system, participants had more menus and menu options to negotiate than when using the simple version. In each version, participants used voice commands to work through the menus and menu options. The findings of Lee et al. (2001) point to an important role for task complexity in assessing the effects of speech-based interfaces on subjective workload and perceived distraction. Task complexity was not manipulated in the current study, but would be an important factor to consider in any future investigations and also in discussing the potential practical implications of the present findings.

It is worth relating the subjective workload results to the results from the objective data (reported in detail elsewhere – Mitsopoulos-Rubens et al., 2013). In brief, analysis of the objective performance data revealed that interacting with a Visual-Manual interface, but not a Visual-Voice or Audio-Voice interface, while driving adversely affected participants’ ability to maintain a constant lane position relative to Baseline driving. The Visual-Manual interface was also associated with significantly more eyes off road time than either of the two voice-based interface types, and significantly more long, safety-critical glances (i.e., greater than 2 seconds; Klauer et al., 2006) to the display than the Audio-Voice condition. Performance on the licence number task was also worst when participants used the Visual-Manual interface. Further, while task completion time did not differ significantly between the Visual-Voice and Audio-Voice interface types, accuracy did, with participants making significantly fewer errors when using the Visual-Voice interface than when using the Audio-Voice interface. Given the subjective workload findings there appears to be some indication that the participants were aware of the added demands associated with performing a secondary task while driving, irrespective of the interface being used. However, while participants were able to compensate and adapt to the added task demands when using the voice-based interface types, the nature of the Visual-Manual interface was such that participants could not compensate and adapt to the same degree.

In summary, these results suggest that the use of a Visual-Manual interface while driving is problematic and so should be avoided by drivers of police vehicles. Given that, in practice, this may be difficult to achieve, the development of voice-based interfaces for in-vehicle police work is encouraged. In saying this it is important to highlight that the use of voice-based interfaces does not eradicate the potential for driver distraction. Thus, in designing a voice-based interface for use by drivers of police vehicles, consideration needs to be given to the full range, nature and complexity of the tasks which police officers are required to perform using the MDT. Not only will this help to ensure that the interface is appropriately acceptable, but critically, will aim to ensure that the interface is as minimally distracting as is possible and the least likely to undermine safety.

Acknowledgements

This research was funded by WorkCover NSW and was conducted in collaboration with the Police Association of NSW (PANSW) and NSW Police. From PANSW, we would like to thank Mr Greg Chilvers and Mr Paul Hannen. From NSW Police, we thank, in particular, Inspectors Peter Hinds and Sean Rogers. From MUARC, we would like to thank Mr Christian Ancora, Mr Kenny Dao, Ms Casey Mackay, Mr Nebojsa Tomasevic, and Mr Glenn Watson. The views expressed here are those of the authors and are not necessarily those of WorkCover NSW, PANSW or NSW Police.

References

Byers, J.C., Bittner, A.C., Hill, S.G. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A, Mital (Ed.), *Proceedings of the Annual International Industrial Ergonomics and Safety Conference: Advances in Industrial Ergonomics and Safety I* (pp. 481-485). London: Taylor & Francis.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*, 319-340.

Garay-Vega, L., Pradhan, A.K., Weinberg, G., Schmidt-Nielsen, B., Harsham, B., Shen, Y., Divekar, G., Romoser, M., Knodler, M., & Fisher, D.L. (2010). Evaluation of different speech and touch interfaces to in-vehicle music retrieval systems. *Accident Analysis and Prevention*, *42*, 913-920.

Hampton, P., & Langham, M. A contextual study of police car telematics: The future of in-car information systems. *Ergonomics*, *48*, 109-118.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., & Ramsey, D.J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (Report No. DOT HS 810 594). National Highway Safety Administration, Washington, DC.

Lee, J.D., Caven, B., Haake, S., & Brown, T.L. (2001). Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors*, *43*, 631-640.

Maciej, J., & Vollrath, M. (2009). Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accident Analysis and Prevention*, *41*, 924-930.

Mitsopoulos-Rubens, E., Filtness, A.J., & Lenné, M.G. (2013). *Injuries to police officers associated with police vehicle travel – phase 2. Report 5: Simulator study.* Draft report prepared for WorkCover NSW through the Police Association of NSW, 5 April.

Mitsopoulos-Rubens, E., Trotter, M.J., & Lenné, M.G. (2011). Effects on driving performance of interacting with an in-vehicle music player: A comparison of three interface layout concepts for information presentation. *Applied Ergonomics*, *42*, 583-591.

Mitsopoulos-Rubens, E., Trotter, M.J., & Rudin-Brown, C.M. (2009). *The interaction between police vehicles and police equipment and its role in injuries to police officers. Report 3: Survey.* Final report prepared for WorkCover NSW through the Police Association of NSW, 2 July.

Ranney, T.A., Harbluk, J.L., & Noy, Y.I. (2006). Effects of voice technology on test track driving performance: Implications for driver distraction. *Human Factors*, *47*, 439-454.

Young, K.L., Lenné, M.G., Archer, J., & Williamson, A.R. (2009). Development and validation of a naturalistic driver distraction evaluation tool. *Proceedings of the First International Conference on Driver Distraction and Inattention.* Gothenburg, Sweden.

Young, K.L., Mitsopoulos-Rubens, E., Rudin-Brown, C.M., & Lenné, M.G. (2012). The effects of using a portable music player on simulated driving performance and task-sharing strategies. *Applied Ergonomics*, *43*, 738-746.