

Unimodal and Multimodal Signals to Support Control Transitions in Semiautonomous Vehicles

Katri Salminen^{1,2}, Ahmed Farooq¹, Jussi Rantala¹, Veikko Surakka¹, Roope Raisamo¹

¹Tampere Unit for Computer-Human Interaction
University of Tampere, Finland

²Tampere University of Applied Sciences, Finland

{katri.salminen, ahmed.farooq, jussi.rantala, veikko.surakka, roope.raisamo}@tuni.fi

ABSTRACT

Semiautonomous driving still requires the driver's control and attention in certain situations. Especially control transitions, i.e. take-over and hand-over situations, are important for safety. Our aim was to study control transitions supported by unimodal (i.e. visual, auditory, or haptic) or multimodal (i.e. visual, auditory and haptic) signals indicating change from manual to autonomous driving and vice versa. The signals were abstract visual blinks, auditory beeps, or haptic vibrations. The task was to take over driving while either looking through the windshield or playing a game. In addition, in half of the control transitions a feedback signal indicated successful control transition. The results showed that a secondary task slowed down the reaction times, but there was a great variation between individuals. In general, the response to auditory signal was slower than to visual, haptic, or multimodal signals. Moreover, users preferred feedback during control transitions but this slowed down the reaction time.

Author Keywords

control transitions, multimodality, human factors, user study

CCS Concepts

• Human-centered computing → Interaction design → Interaction design process and methods → Activity centered design

INTRODUCTION

Self-driving cars are emerging in the market. In practice, during automated driving, the driver's *primary task* shifts from driving to monitoring. Unfortunately, high levels of automation can have detrimental effects [22, 43, 44]. While the fully autonomous mode of driving is on, drivers will focus on *secondary tasks* like gaming or conversations with passengers [6]. It can take over 20 seconds for the driver to disengage from secondary task and to be functional for the required driving-related actions like braking [24]. Further, autonomous cars still suffer from frequent system failures

and therefore require driver's assistance [10]. Taken together, there is a need to make drivers aware of the traffic, limitations of the system, and when to take control over the driving [7].

Some earlier work has aimed to solve the problem by providing continuous information to the driver during automated driving [2, 5, 46]. However, continuous information mediation about, for example, ongoing traffic or the functionality of the automated driving system can be distracting or neglected. Drivers easily miss dynamic and directional cues [34], complex cues do not necessarily prepare the driver for a correct action [33], and environmental issues like stereo volume has to be taken into account when mediating detailed information [14]. Therefore, most previous studies have concentrated on control transitions, in this paper referred as hand-over and take-over requests (i.e. HOR and TOR). By HOR we refer to a signal indicating that the automated driving mode is on and the control from manual driving can be released to the system. By TOR we refer to a signal indicating the need to take over driving tasks. Such control transition requests can be spoken or text-based messages, abstract unimodal (mostly visual, auditory, or haptic and sometimes even smell) or multimodal cues [1, 11, 13, 16, 29, 36, 37, 38, 41].

One important aspect related to successful take-over is correct timing. Several studies suggest that at least five seconds is needed for the driver to avoid hazard (e.g. an obstacle in the road) safely [12, 30, 31]. It can be noted [19] that not even five seconds may sufficient for avoiding hazard. Longer times (e.g. 7 seconds) improve take-over quality and errors as drivers can use the extra time for decision making. However, often it is not possible to present TOR in advance, and therefore the majority of previous studies has measured reaction times after a TOR. A recent overview [13] on the topic showed that on average TOR reaction times varied between studies from 1.14 seconds to up to 15 seconds. The variation in reaction times

is due to several reasons. For example, driver's reactions and preferences are depending on the nature of control transition requests, secondary task modality (e.g. speaking or browsing a mobile phone), and the experimental design. Previous studies show that on average, engaging any kind of secondary task significantly slows down take-over times [13] and that tasks requiring visual attention result in significant amounts of driving errors [49]. In addition, visual TOR may not evoke as fast reactions as auditory, haptic, or multimodal ones [4, 35, 39, 50]. For example, in a previous study [38] visual text-based TOR took on average 6.9 seconds to be noticed while auditory, tactile, and multimodal TOR required on average less than 3 seconds to be noticed. Third, there might be some tendency to react faster to multimodal than unimodal TOR [39, 47]. Petermeijer et al. [33] showed that drivers reacted to audio-tactile TOR faster than to auditory only or tactile only requests. The differences, however, were present only when the participants touched the steering wheel and diminished, when the task was, for example, to break after TOR.

Taken together, due to varying methodology affecting results of the studies, there is still need and room to study control transitions further. We studied reaction times, driving behavior, and driver preferences to abstract unimodal (i.e. blinking lights, auditory beeps and seat vibrations) and multimodal (i.e. combination of those three) HORs and TORs. We had following objectives in the study: a) To study potential differences between modalities if HORs and TORs are designed in a manner that they should be noticeable even when the driver is playing with a mobile phone. b) To study if feedback indicating successful take-over can improve user's preferences regarding the system and accelerate driver's reactions. c) Describe the technical setup in detail using standardized and tested methods so that the results are valid and repeatable.

METHODS

The technical setup consisted of three areas: the driving simulator, the type of request (visual, auditory and haptic), and simulation of manual and autonomous driving modes. These are described below.

Driving simulator setup and logging software

A fixed-base driving simulator was implemented in a laboratory setup (see [15, 32] for examples of a fixed-base simulators). A PC ran the lane change test (LCT) driving simulator software (ISO 26022, 2010). We selected the LCT over other simulators to ensure a controlled testing environment with regards to road traffic, geographical landmarks and other obstacles. The participants were

instructed to drive at a constant speed of 100km/h, during which the software provided visual instructions every five seconds of how to navigate. The task was to navigate to the instructed lane as quickly as possible. The simulator was run on a 55 inch LCD display (Samsung UNES7005) brightness set to 28lux with an ambient room brightness of 7lux within the laboratory.

The custom logging software followed the client/server model. The client application was run on the same PC that ran the LCT simulator. The client application logged all user inputs to the system using a custom key-logger, that was linked to the Logitech G27 steering wheel and pedals. All user activity including the steering wheel movements (i.e. steering wheel reversals and deviation from central line), the buttons on the steering wheel (for engaging and disengaging semiautonomous mode) along with the various peddles were recorded to within a 5ms accuracy. Furthermore, the LCT inbuilt logger also recorded the user driving behavior using the GMT_ms (GMT time in milliseconds) function. All measurements were averaged over three samples and taken from the current driving position (i.e. the position in the road of the driving simulator). The client end of the system was used as the central source for controlling the three signals (i.e. audio, visual, and haptic). This was done using a USB to serial FTDI 232R interface that controlled the visual signal mechanisms. The haptic feedback was also triggered using the FTDI 232R interface, but this was connected to a custom signal generator and D-class amplifier with a pre-calibrated sine wave actuation signal (as detailed in the following section). Audio was generated through the ASCII bell function within the java client using a separate thread. This meant that all three modalities and the logging software were allocated their own threads and worked together at the client end to provide signal and logged every system / user generated event in real-time to a high degree of precision.

The server application was run on a separate PC to ensure no bottlenecks were created on the client machine. The server was used to log eternal events and to manage the switching of the different display devices (see Fig. 1). The server also generated the HORs and TORs along with the ability to create multimodal signal for the user, remotely. This was done to ensure that transitions between autonomous and manual mode were logged-in properly and that the experimenter could manually generate these events at various stages of the LCT track without any logging issues.

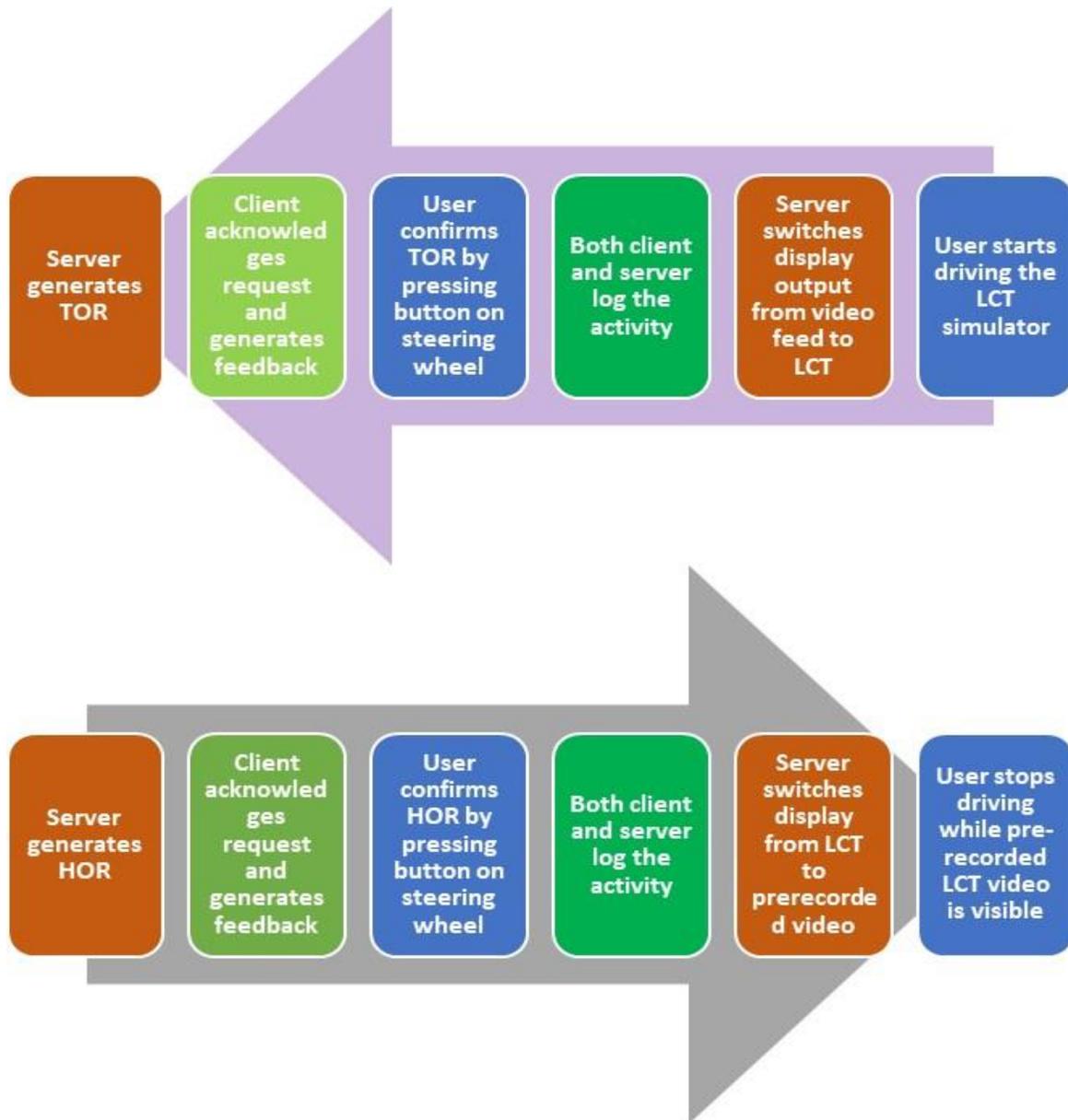


Fig. 1. Sequence of events involved during TO (top) and HO (bottom).

Visual, haptic, and audio implementation

Visual signal was provided through a custom HUD implemented in accordance with, e.g., [25]. The design was modeled after the Volvo XC90 City Safety System, which used driver notification alert by light emitting diodes (LEDs) at the front of the dashboard. Our implementation consisted of a series of 23 red LEDs (WW05A3SRP4-N2) spaced out over a total length of 172 mm. The duration of visual alerts was fixed at 900ms. The signal was broken down into two chunks, each of 400ms, with a delay of 100ms between them. Each LED had a relative luminance of 1.5 a.u. at 30 mA. This

was used as the base brightness parameter for the visual signal with a combined brightness of 45 lux. The LED strip was hidden behind a cover so that only the light reflected via windscreen was visible for the participant. The display brightness of 28 lux was not considered to have hindered the HUD as the HUD was about twice as bright. The HUD was also clearly distinguishable from the LCT display, as the reflection of the HUD from the display (Samsung UNES7005) was very prominent.

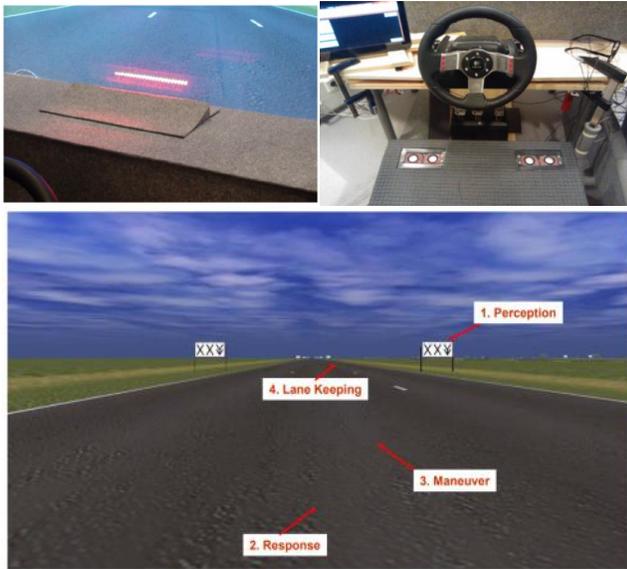


Fig. 2. The visual (left above) and haptic (right above) interaction modalities used to generate HOR and TOR while using the LCT simulation software (below).

The haptic signal was provided through a custom implementation of the Haptic Seat (see Fig. 2, and [32] or [45] for more information about such seats). The current implementation provided haptics with integrated actuators located under the driver's thighs. The prototype utilized two Tectonic TEAX25C10-8/HS voice coil actuators embedded into the lower part of the driver's seat. To increase the signal area, these actuators were fitted with a horizontal plastic extension of 6 inches in length. The Tectonic actuators were connected to an amplifier and a signal generator. The signal generator was set to provide a sinusoidal wave at 150 Hz with a peak amplitude value of 18.8V. The length of the haptic signal was fixed at 500 ms so that it consisted of two 200 ms bursts with 100 ms pause between them.

Audio was provided through Samsung UNES7005 television's internal speakers, where the amplitude of the audio was kept between 60 – 68 dB, irrespective of the source device. This included the audio relevant to the LCT simulator as well as the audio HO / TO request. The LCT ambient engine noise was between 50-51dB while the maximum engine noise while driving was between 58-60dB. The audio stimulus was structured similarly to the haptic stimulus and consisted of two ASCII bell outputs generated through a Java client application. The length of the auditory signal was fixed at 500ms so that it consisted of two 200 ms long beeps with 100 ms pause between them. The total length of the entire stimulus was 500 ms with a delay of 100ms between each segment. The audio was measured at 68db and was clearly audible over the engine noise of the LCT.

Simulating manual and autonomous modes

As it is not possible to simulate autonomous driving mode in the LCT software, our setup utilized pre-recorded sessions to achieve this. The sessions were timeline marked, which meant that it was possible to navigate to any specific section of the recording to match the user's position on the track. As illustrated by Fig. 3, after HO request the user display input was switched to the pre-recorded driving session video, specifically to the exact section of the track corresponding to the current user lane position on the LCT software. The switching was done using a Roland V-1HD device, which seamlessly transitioned from source 'A' (LCT simulator) to source 'B' (pre-recorded LCT video session). Although the switcher was capable of automatic trigger-switching using the serial port input, manual transition was preferred to ensure completely seamless transition. Take-over transitions were done in a similar manner. Once the user confirmed the TO request by pressing the relevant button on the steering wheel, the Roland V-1HD device was used to switch the display input



from the pre-recorded LCT video session to the LCT simulator software. The driver was instructed to press the accelerator pedal so that the vehicle was moving once the user took control over the driving.



Fig. 3. The driving setup (top) and LCT simulator with overlaid turn instructions (bottom).

One of the conditions in the experiment was to have the user perform a specific task after TO request. As the LCT simulation software does not have supporting mode, we utilized input merging as well as switching options on the Roland V-1HD device. To achieve this, we used a third PC input with "turn left" and "turn right" markers and merged this input with the LCT simulator visual input (Fig. 3). This meant that the user display would have the LCT simulator input overlaid with either the "turn left" or the "turn right"

markers. Using this method, we were able to instruct the user to perform a specific turning task. External events were included in the logs and all input transitions were registered in the server_log file with a date_time stamp. Using this technique, it was possible to measure the exact time when the task was introduced and the time it took to move to the specific lane by the user by comparing the server, client, and LCT logs.

Participants

A total of 12 participants (3 females) took part in the study. Their average age was 31 years, range 18 – 57. The average time since obtaining the driver's license was 13 years (range 1 – 38), and on average the participants drove 5800 kilometers per year (range 3000 – 25 000). None of the participants had previous experience with autonomous driving. All were students or staff from the Tampere University.

Procedure

The experiment utilized two designs. A 4×3 (experimental block \times request modality) design was used to test control transitions with unimodal signals and a 2×2 (feedback \times direction) to test them with multimodal ones. The procedure for unimodal condition was the following. The participants were allowed to practice the procedure and requests since it has been shown to improve take-over performance [21]. First, the participant was familiarized with the manual driving task with LCT simulator. When the driving was fluent, the first HOR was presented. The participant was told to press a button on a steering wheel to indicate notification of the request, and then release hands. During the autonomous driving, the task was to monitor driving through the windshield. Then, a TOR was presented. The task was to take over driving by grabbing the steering wheel and pressing a button indicating successful take over as quickly as possible, and then immediately begin manual driving according to LCT procedure. This was repeated until visual, auditory, and haptic requests were introduced. Only one modality was used at a time so that, for example, visual HOR was always paired with visual TOR. Before the experiment began, the participant was able to try the mobile game Rise Up. In the game, the player is protecting an air balloon by moving objects ahead able to burst the balloon.

In unimodal condition, there were four experimental blocks. The level of distraction and use of feedback was varied between the blocks. In two blocks the participant was asked to monitor autonomous driving through the windshield and in two blocks to play a mobile phone a game. Also, in two blocks there was no feedback for confirming the transition on and in two blocks an auditory beep was played to indicate that the control transition had taken place. The order of the blocks was counterbalanced. Each block had a total of 3 trials presented in random order. An experimental block proceeded similarly than familiarization. After 120 to 240 seconds manual driving, a HOR was initiated. A TOR was launched after 120 to 240 seconds of autonomous

driving. After completing one block, a questionnaire was presented. The questions were “would you use the TORs system?” and “which modality was most effective in getting your attention?”. After completing all the four blocks, the participants filled a questionnaire with following questions: “did you prefer feedback indicating successful hand-overs and take-overs”, “which modality did you prefer (visual, auditory, or haptic)?”, “which modality was most effective?”, and “which modality did you find most reliable?”.

The multimodal condition was always the last block. It used similar control transition procedure as unimodal condition, but the control transition request combined all the modalities (i.e. visual, audio, and haptic). The block consisted of four individual tasks. The task was either to turn right or left after TOR as quickly and accurately as possible based on visual command on the screen. The participants were instructed to always play the mobile phone game during autonomous driving. Two of the tasks had feedback and two did not. Then, the participants were asked to fill a questionnaire. The questions were “did multimodality improve the clarity of the request?” and “did feedback about successful take-over make turning right or left easier?”. A total amount of experimental trials was 16. Conducting a session in the experiment took a total of 45 min.

Data Analysis

A two way within-subjects 4×3 (experimental block \times request modality) repeated measures analysis of variance (ANOVA) was conducted to analyze reaction times for HOR and TOR. HO reaction time was the time from the end of the HOR to the time the participants successfully pressed the steering wheel button to initiate hand-over. TO reaction time was measured starting from the TOR till the participants took control of the steering wheel and disengaged the semi-autonomous mode by pressing the button of the steering wheel. Pairwise t-tests with Bonferroni corrected p-values were used for post hoc tests. Due to the different design, the reaction time data from multimodal blocks was excluded from the comparison between the four other blocks. However, in the results section the means and standard errors of the means (SEMs) are presented for comparison.

In addition, a two way 2×2 (feedback \times direction) ANOVA was performed for multimodal block's TOR. Further, the task completion time was tested using a 2×2 (feedback \times direction) ANOVA. Two measures were used to calculate the task completion: the time to initiate the movement after the task was visually given to the participant, and time from task visibility to achieving steady driving on the median point of a lane according to LCT data. Partial eta squared (partial η^2) are reported after all the ANOVAs.

RESULTS

Reaction times to HORs

A 4×3 ANOVA showed no statistically significant main or interaction of main effects for HO reaction times (see Table 1). Partial η^2 were 0.05 for block, 0.06 for modality, and 0.09 for block \times modality. Table 2 shows similar numbers for multimodal HORs. Partial η^2 were 0.16 for block, 0.13 for modality, and 0.05 for block \times modality.

		M	Sd	Range
Block 1: No mobile phone, no feedback	V	1,14	0,47	0,56 – 2,05
	A	1,22	0,52	0,44 – 2,26
	H	1,36	0,68	0,55 – 3,16
Block 2: No mobile phone, feedback	V	1,19	0,36	0,60 – 1,79
	A	1,34	0,65	0,76 – 3,08
	H	1,12	0,14	0,78 – 1,31
Block 3: Mobile phone, no feedback	V	1,27	0,50	0,63 – 2,42
	A	1,15	0,47	0,74 – 2,29
	H	1,22	0,40	0,64 – 1,77
Block 4: Mobile phone, feedback	V	1,21	0,40	0,75 – 2,00
	A	1,45	0,37	0,92 – 2,12
	H	1,22	0,52	0,63 – 2,21

Table 1. Means (M), standard deviations (Sd) and ranges (min to max) for visual (V), auditory (A), and haptic (H) HORs in each four experimental blocks.

	M	Sd	Range
No feedback, left	1,54	0,81	0,57 – 2,73
No feedback, right	1,19	0,47	0,37 – 1,84
Feedback, left	1,21	0,67	0,43 – 2,73
Feedback right	1,10	0,57	0,58 – 2,73

Table 2. Means, standard deviations and ranges for multimodal HORs in each four tasks.

Reaction times to TORs

A 4×3 ANOVA showed statistically significant main effects for experimental block $F(3, 33) = 25.7$, $p < 0.001$, partial $\eta^2 = 0.7$ and modality $F(2, 22) = 4.6$, $p < 0.05$, partial $\eta^2 = 0.30$ (see Table 3). The interaction of the main effects was not statistically significant, partial $\eta^2 = 0.13$. Post hoc comparisons showed that the participants reacted

faster to TOR in block 1 than in block 3 (mean difference (md) = 0.8, $p < 0.001$) and 4 (md = 1.0, $p < 0.001$), and also faster in block 2 than in block 3 (md = 0.7, $p < 0.01$) or 4 (md = 0.9, $p < 0.01$). Post hoc comparisons showed no statistically significant differences between modalities. However, the difference between haptic and auditory TOR was approaching significance (md = 0.12, $p = 0.057$) suggesting that at least in some cases haptic TOR can yield faster reaction times.

		M	Sd	Range
Block 1: No mobile phone, no feedback	V	1,34	0,35	0,85 – 2,09
	A	1,29	0,22	0,91 – 1,74
	H	1,24	0,30	0,78 – 1,87
Block 2: No mobile phone, feedback	V	1,38	0,26	1,09 – 1,81
	A	1,47	0,21	1,16 – 1,82
	H	1,32	0,25	1,01 – 1,83
Block 3: Mobile phone, no feedback	V	1,99	0,52	1,01 – 2,89
	A	2,26	0,61	1,26 – 3,21
	H	2,08	0,59	1,15 – 3,37
Block 4: Mobile phone, feedback	V	2,42	0,89	0,93 – 4,11
	A	2,31	0,71	1,19 – 3,22
	H	2,20	0,74	0,83 – 3,11

Table 3. Means, standard deviations and ranges (min to max) for visual, auditory, and haptic TORs in each four experimental blocks

A 2×2 (feedback \times direction) ANOVA for multimodal TORs showed a statistically significant main effect for feedback $F(1, 11) = 8.4$, $p < 0.05$, partial $\eta^2 = 0.43$ (see Table 4). The main effect for direction or the interaction of the main effects were not statistically significant, partial $\eta^2 = 0.003$ and 0.004, respectively. Post hoc comparisons showed that the participants reacted faster to TORs without than with feedback (md = 0.32, $p < 0.05$).

	M	Sd	Range
No feedback, left	1,80	0,51	1,05 – 2,50
No feedback, right	1,80	0,45	1,14 – 2,64
Feedback, left	2,13	0,60	1,19 – 2,97

Feedback, right	2,11	0,61	1,11 – 3,15
-----------------	------	------	-------------

Table 4. Means, standard deviations and ranges for multimodal TORs in each four tasks.

Task completion after multimodal TOR

A 2 × 2 ANOVA showed no statistically significant main effects (partial $\eta^2 = 0.08$ for feedback and 0.15 for direction) or interaction of the main effects (partial $\eta^2 = 0.27$) in movement times (see Table 5 and 6). A 2 × 2 ANOVA showed a statistically significant main effect of direction $F(1, 11) = 4.88, p < 0.05, \text{partial } \eta^2 = 0.31$ for the time for task visibility to achieving median point of the lane. Post hoc comparisons showed that the participants completed turning left faster than turning right ($md = 0.27, p < 0.05$). The main effect for feedback or the interaction of the main effects were not statistically significant, partial $\eta^2 = 0.03$ and 0.05, respectively.

	M	Sd	Range
No feedback, left	0,64	0,17	0,34 – 0,93
No feedback, right	0,59	0,20	0,26 – 0,95
Feedback, left	0,58	0,29	0,26 – 1,25
Feedback, right	0,95	0,82	0,40 – 3,47

Table 5. Means, standard deviations and ranges for movement times in each four tasks.

	M	Sd	Range
No feedback, left	3,04	1,15	1,95 – 5,85
No feedback, right	3,14	0,84	1,87 – 4,77
Feedback, left	2,98	0,87	2,12 – 5,01
Feedback, right	3,43	1,43	2,06 – 6,46

Table 6. Means, standard deviations and ranges for task completion times in each four tasks.

Questionnaires

Questions about preferences revealed that all the twelve participants would use unimodal (i.e. visual, audio or haptic) HORs and TORs. Nine participants considered visual signals the most effective while not playing, two audio, and one the haptic signal. However, when the

participants were playing, feedback affected the responses. When there was no feedback, six participants preferred the audio, five the visual, and one preferred the vibration signal. When there was feedback, five participants preferred audio, four light, and three vibration. After conducting all the four unimodal blocks, ten participants preferred feedback. However, only five participants considered that feedback was aiding in completion of the driving task in multimodal block. Six participants rated light, four audio, and two vibration as most pleasant. Seven participants rated light, four audio, and one vibration as most effective. Finally, six participants rated light, four audio, and two vibration as most trustworthy. Seven participants considered that multimodal HORs and TORs improved the trustworthiness of the system when compared to the unimodal ones.

DISCUSSION

The results provided valuable information about the timing, functionality, and preference of the control transitions using varying modalities, and the use of feedback related to successful transition. The reaction times with all unimodal and multimodal HORs were on average less than 1.4 seconds. Some studies have found similar reaction times [13], while especially in the case of language-based warnings reaction times have at times been significantly longer [38]. In line with previous studies, our results suggest that abstract signals result to faster reaction times for HOR than language-based ones. Further, the current results suggest that modality has no effects to the HO reaction time.

The reaction times to TORs were comparable to HORs when the participant’s task was to monitor driving. The modality had no effect on the reaction time in this case. However, playing a mobile phone slowed down reaction times as expected [13]. Since the participants were instructed to detach their hands from steering wheel also while not playing, the hand movements are unlikely to be the reason behind the result. On average, TO times were 2.19 seconds, less than one second slower than reactions to HORs. In comparison to previous studies [13, 38] the reaction times were relatively fast (for example, see [13] reporting TO reaction times varying from 2 to 3.5 seconds or even more). The longest reaction times are often due to requests requiring information processing prior to reaction [38]. Further, it is quite common to miss the TORs altogether [13]. In the current study, only one participant had problems noticing visual TORs in one case while playing with a mobile phone. Thus, it seems that overall even visual abstract TORs functioned well in respect to reaction times when compared with previous research. This supports the design choices made in implementation of the conditions. It should, however, be noted that investigation of the data shows that while some participants reacted to the TORs quickly while playing (e.g. in less than a second after the request), there were participants who took more than 4 seconds to react to the request. This result can be observed

from the light of previous research showing that the urgency of the request needs to be mediated to the driver if a fast reaction is required [39]. It may be that the slowest participants simply prioritized playing over the take-over.

The TOR modality had some effect on reaction times. On average, the participants reacted to haptic and visual TORs faster than to auditory ones. Due to the small sample size the result was not statistically significant, even though the difference in average reaction times was close to 0.5 seconds. Most previous studies [17, 35, 39, 50] have suggested that visual TORs take more time to respond than auditory or haptic ones, especially if the driver performs a secondary task. In our study, this was not the case. This is likely due to the implementation of the visual request so that the LED displays were stimulating peripheral vision [see, for example, 3, 23, 26, 27, 28 for additional information about LED displays inside vehicles]. In addition, multimodal TOR was on average reacted faster than unimodal ones, which is in line with the previous studies [36, 47]. However, the effect was not present with the HORs meaning that it might be that the multimodal TORs are most effective only while performing a secondary task.

The questionnaire showed that the least preferred modality was always haptics. These results contradict previous findings related to haptic seats [8, 45] as well as reaction times. In the case of haptics, it is possible, that the participants had difficulties to associate seat vibration to the reaction needed. This was at least partly supported by post-experimental comments by the participants, in which two participants reported dislike for haptic request implemented as it was done. By using, for example, thermal or shape-changing haptics instead of vibrotactile stimulation it might be possible to alter driver's preferences and performance [9, 20]. The most preferred modality was visual, even though there was a small shift in preference towards auditory modality while playing with the mobile phone. This result indicates that a LED display is functional for a control transition request even while performing a secondary task. Finally, most participants did consider multimodal HORs and TORs as the most effective. A reason behind the result

might be that by stimulating more modalities the level of activity is elevated [42]. Together, the results indicate that the implementation of a control transition request should be considered carefully so that, for example, visual requests can be seen while performing secondary tasks and that haptics are pleasant.

Finally, unimodal feedback for confirming the transfer was preferred. However, the participants did not consider it useful in performing driving-related tasks, and in general, task completion times with the feedback were slower than without it. The result seems to confirm previous findings suggesting that effects of confirmation feedback on driving performance are not consistent [40]. As the confirmation feedback in the current study was always an auditory beep, it should be noted that feedback directing the attention of the driver to the task and its completion (e.g. blinking lights at right or left) might have yielded different results. Environmentally responsive systems sensitive to, for example, urgency of action [18, 48] are in the scope of our longer-term research.

CONCLUSIONS

The results showed that control transitions can be performed by using abstract and simple hand-over and take-over requests with any of the given modalities. The reactions to the auditory signal were slowest ones. The reason behind this might be that the design of a signal affects to the effectiveness of the modality. Performing a secondary task slowed down the reaction times, but there was a great variation between individuals. This result suggests that providing information related to transition urgency or situation around the vehicle could speed up the reaction times as the participants would not concentrate on gaming. Finally, while the feedback during control transitions was preferred, it also slowed down the reaction times. To find out reasons behind this result, further experiments focusing solely on the use of feedback in varying driving and secondary tasks are required. Our next study will move from highly controlled laboratory experiment to a study with a more realistic driving simulator that adds to the complexity of events and information received.

ACKNOWLEDGMENTS

This work was supported by Henry Ford Foundation (grant 201800029) and project Multimodal In-Vehicle Interaction and Intelligent Information Presentation (MIVI), funded by Business Finland (grant 8004/31/2018).

REFERENCES

1. Bazilinskyy, P., Petermeijer, S. M., Petrovych, V., Dodou, D., & de Winter, J. C. (2018). Take-over requests in highly automated driving: A crowdsourcing survey on auditory, vibrotactile, and visual displays.

Transportation research part F: traffic psychology and behaviour, 56, 82-98.

2. Beattie, D., Baillie, L., Halvey, M., & McCall, R. (2014, October). What's around the corner?: enhancing driver awareness in autonomous vehicles via in-vehicle spatial auditory displays. In *Proceedings of the 8th nordic conference on human-computer interaction: fun, fast, foundational* (pp. 189-198). ACM.
3. Borojeni, S. S., Chuang, L., Heuten, W., & Boll, S. (2016, October). Assisting drivers with ambient take-over requests in highly automated driving. In *Proceedings of the 8th International Conference on*

Automotive User Interfaces and Interactive Vehicular Applications (pp. 237-244). ACM.

4. Borojeni, S. S., Wallbaum, T., Heuten, W., Boll, S. (2017). Comparing Shape-Changing and Vibro-Tactile Steering Wheels for Take-Over Requests in Highly Automated Driving. In proceedings of the 9th ACM International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '17), <http://dx.doi.org/10.1145/3122986.3123003>
5. Cohen-Lazry, G., Borowsky, A., & Oron-Gilad, T. (2017, September). The effects of continuous driving-related feedback on drivers' response to automation failures. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 61, No. 1, pp. 1980-1984). Sage CA: Los Angeles, CA: SAGE Publications.
6. Clark, H., & Feng, J. (2017). Age differences in the takeover of vehicle control and engagement in non-driving-related activities in simulated driving with conditional automation. *Accident Analysis & Prevention*, *106*, 468-479.
7. Cunningham, M., & Regan, M. A. (2015, October). Autonomous vehicles: human factors issues and future research. In *Proceedings of the 2015 Australasian Road safety conference* (pp. 14-16).
8. Dass Jr, D. E., Uyttendaele, A., & Terken, J. (2013, October). Haptic in-seat feedback for lane departure warning. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 258-261). ACM.
9. Di Campli San Vito, P., Brewster, S., & Pollick, F. (2018). Investigation of Thermal Stimuli for Lane Changes.
10. Dikmen, M., & Burns, C. M. (2016, October). Autonomous driving in the real world: Experiences with tesla autopilot and summon. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 225-228). ACM.
11. Dmitrenko, D., Maggioni, E., Vi, C. T., & Obrist, M. (2017, September). What did i sniff?: Mapping scents onto driving-related messages. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 154-163). ACM.
12. Epple, S., Roche, F., & Brandenburg, S. (2018, September). The sooner the better: Drivers' reactions to two-step take-over requests in highly automated driving. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 62, No. 1, pp. 1883-1887). Sage CA: Los Angeles, CA: SAGE Publications.
13. Eriksson, A., & Stanton, N. A. (2017). Takeover time in highly automated vehicles: noncritical transitions to and from manual control. *Human factors*, *59*(4), 689-705.
14. Fagerlönn, J., Lindberg, S., & Sirkka, A. (2012, October). Graded auditory warnings during in-vehicle use: using sound to guide drivers without additional noise. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 85-91). ACM.
15. Farooq A., Evreinov G., Raisamo R., Hippula A. (2019). Developing Intelligent Multimodal IVI Systems to Reduce Driver Distraction. In Proceedings of the International Conference on Intelligent Human Systems Integration (IHSI), Feb 7-10, 2019, San Diego, California.
16. Forster, Y., Naujoks, F., Neukum, A., & Huestegge, L. (2017). Driver compliance to take-over requests with different auditory outputs in conditional automation. *Accident Analysis & Prevention*, *109*, 18-28.
17. Gable, T. M., Walker, B. N., Moses, H. R., & Chitloor, R. D. (2013, October). Advanced auditory cues on mobile phones help keep drivers' eyes on the road. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 66-73). ACM.
18. Gaspar, J. G., Brown, T. L., & Marshall, D. C. (2015, September). Examining the interaction between timing and modality in forward collision warnings. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 313-319). ACM.
19. Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013, September). "Take over!" How long does it take to get the driver back into the loop?. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 57, No. 1, pp. 1938-1942). Sage CA: Los Angeles, CA: SAGE Publications.
20. Grah, T., Epp, F., Wuchse, M., Meschtscherjakov, A., Gabler, F., Steinmetz, A., & Tscheligi, M. (2015, September). Dorsal haptic display: a shape-changing car seat for sensory augmentation of rear obstacles. In Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 305-312). ACM.
21. Hergeth, S., Lorenz, L., & Krems, J. F. (2017). Prior familiarization with takeover requests affects drivers' takeover performance and automation trust. *Human factors*, *59*(3), 457-470.
22. Kerschbaum, P., Lorenz, L., Hergeth, S., & Bengler, K. (2015, June). Designing the human-machine interface for highly automated cars—Challenges, exemplary concepts and studies. In *2015 IEEE International*

Workshop on Advanced Robotics and its Social Impacts (ARSO) (pp. 1-6). IEEE.

23. Langlois, S. (2013, October). ADAS HMI using peripheral vision. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 74-81). ACM.
24. Lu, Z., Coster, X., & de Winter, J. (2017). How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving. *Applied ergonomics*, *60*, 293-304.
25. Lylykangas, J., Surakka, V., Salminen, K., Farooq, A., & Raisamo, R. (2016). Responses to visual, tactile and visual-tactile forward collision warnings while gaze on and off the road. *Transportation research part F: traffic psychology and behaviour*, *40*, 68-77.
26. Löcken, A., Heuten, W., & Boll, S. (2015, September). Supporting lane change decisions with ambient light. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 204-211). ACM.
27. Löcken, A., Heuten, W., & Boll, S. (2016, October). Enlightening drivers: A survey on in-vehicle light displays. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 97-104). ACM.
28. Meschtscherjakov, A., Döttlinger, C., Rödel, C., & Tscheligi, M. (2015, September). ChaseLight: ambient LED stripes to control driving speed. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 212-219). ACM.
29. Mirnig, A. G., Gärtner, M., Laminger, A., Meschtscherjakov, A., Trösterer, S., Tscheligi, M., ... & McGee, F. (2017, September). Control Transition Interfaces in Semiautonomous Vehicles: A Categorization Framework and Literature Analysis. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 209-220). ACM.
30. Mok, B. K. J., Johns, M., Lee, K. J., Ive, H. P., Miller, D., & Ju, W. (2015, June). Timing of unstructured transitions of control in automated driving. In *Intelligent Vehicles Symposium (IV), 2015 IEEE* (pp. 1167-1172). IEEE.
31. Mok, B., Johns, M., Lee, K. J., Miller, D., Sirkin, D., Ive, P., & Ju, W. (2015, September). Emergency, automation off: Unstructured transition timing for distracted drivers of automated vehicles. In *2015 IEEE 18th International Conference on Intelligent Transportation Systems* (pp. 2458-2464). IEEE.
32. Nukarinen, T., Rantala, J., Farooq, A., & Raisamo, R. (2015, June). Delivering directional haptic cues through eyeglasses and a seat. In *2015 IEEE World Haptics Conference (WHC)* (pp. 345-350). IEEE.
33. Petermeijer, S., Bazilinskyy, P., Bengler, K., & de Winter, J. (2017). Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied ergonomics*, *62*, 204-215.
34. Petermeijer, S. M., Cieler, S., & De Winter, J. C. (2017). Comparing spatially static and dynamic vibrotactile take-over requests in the driver seat. *Accident Analysis & Prevention*, *99*, 218-227.
35. Petermeijer, S., Doubek, F., & de Winter, J. (2017, October). Driver response times to auditory, visual, and tactile take-over requests: A simulator study with 101 participants. In *Systems, Man, and Cybernetics (SMC), 2017 IEEE International Conference on* (pp. 1505-1510). IEEE.
36. Politis, I., Brewster, S., & Pollick, F. (2013, October). Evaluating multimodal driver displays of varying urgency. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 92-99). ACM.
37. Politis, I., Brewster, S., & Pollick, F. (2014, September). Speech tactons improve speech warnings for drivers. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1-8). ACM.
38. Politis, I., Brewster, S., & Pollick, F. (2015, September). Language-based multimodal displays for the handover of control in autonomous cars. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 3-10). ACM.
39. Politis, I., Brewster, S., & Pollick, F. (2017). Using multimodal displays to signify critical handovers of control to distracted autonomous car drivers. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, *9*(3), 1-16.
40. Roberts, S. C., Horrey, W. J., & Liang, Y. (2012, October). Effect of performance feedback (or lack thereof) on driver calibration. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 67-74). ACM.
41. Roche, F., Somieski, A., & Brandenburg, S. (2018). Behavioral changes to repeated takeovers in highly automated driving: effects of the takeover-request design and the nondriving-related task modality. *Human factors*, 0018720818814963.
42. Salminen, K., Surakka, V., Lylykangas, J., Rantala, J., Ahmaniemi, T., Raisamo, R., ... & Kildal, J. (2012). Tactile modulation of emotional speech samples. *Advances in Human-Computer Interaction, 2012*, 17.

43. Sheridan, T. B., (2002). *Humans and automation: System design and research issues* (Vol. 280). Santa Monica, CA: Human Factors and Ergonomics Society.
44. Strand, N., Nilsson, J., Karlsson, I. M., & Nilsson, L. (2014). Semi-automated versus highly automated driving in critical situations caused by automation failures. *Transportation research part F: traffic psychology and behaviour*, 27, 218-228.
45. Telpaz, A., Rhindress, B., Zelman, I., & Tsimhoni, O. (2015, September). Haptic seat for automated driving: preparing the driver to take control effectively. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 23-30). ACM.
46. van Veen, T., Karjanto, J., & Terken, J. (2017, September). Situation awareness in automated vehicles through proximal peripheral light signals. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 287-292). ACM.
47. Walch, M., Lange, K., Baumann, M., & Weber, M. (2015, September). Autonomous driving: investigating the feasibility of car-driver handover assistance. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 11-18). ACM.
48. Wintersberger, P., Riener, A., Schartmüller, C., Frison, A. K., & Weigl, K. (2018, September). Let Me Finish before I Take Over: Towards Attention Aware Device Integration in Highly Automated Vehicles. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 53-65). ACM.
49. Yang, Y., Reimer, B., Mehler, B., Wong, A., & McDonald, M. (2012, October). Exploring differences in the impact of auditory and visual demands on driver behavior. In *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 173-177). ACM.
50. Yoon, S. H., Kim, Y. W., & Ji, Y. G. (2019). The effects of takeover request modalities on highly automated car control transitions. *Accident Analysis & Prevention*, 123, 150-158.