On-Road Evaluation of In-vehicle Interface Characteristics and Their Effects on Performance of Visual Detection on the Road and Manual Entry

Youngbo Suh, Korea Institute for Defense Analyses, Seoul, South Korea and **Thomas K. Ferris**, Texas A&M University, College Station, USA

Objective: This study investigated the impact of in-vehicle interface characteristics on drivers' multitasking performance measures relating to visual attention management, concerning the distraction potential of in-vehicle touchscreens.

Background: Compared with physical controls providing drivers with naturalistic nonvisual cues, in-vehicle touchscreen interaction relies on vision to a greater extent, leading to more time with eyes off the road and concerns for safety. Little is known from existing research about the extent to which synthetic feedback of in-vehicle touchscreens support visual attention of multitasking drivers, while automakers are increasingly incorporating nondriving functions into in-vehicle touchscreens.

Method: Twenty-nine participants drove an instrumented vehicle on a closed course and acknowledged visual probes obscured on the roadside, while performing a manual data entry task with input interfaces mounted on the center console. The interfaces differed by interface type, key feedback modality, and key size; the configuration of interface characteristics was the within-subject variable. The collected data include performance measures concerning visual detection and touchscreen interaction, in addition to perceived workload.

Results: The addition of nonvisual feedback to touchscreen interaction significantly improved accuracy and promptness of visual detection. No significant difference was found between different sizes of touchscreen keys when synthetic nonvisual feedback was available. Given multisensory feedback, no measure showed a difference between touchscreen conditions and a physical keypad.

Conclusion: The provision of synthetic nonvisual feedback to touchscreen interaction can support visual attention and enhance multitasking performance in driving.

Application: This study can inform in-vehicle interface designers and policy makers concerned with distracted driving and safety.

Keywords: driver attention management, in-vehicle touchscreen, multisensory feedback, multitask performance

Address correspondence to Thomas K. Ferris, Department of Industrial and Systems Engineering, Texas A&M University, 3131 TAMU, College Station, TX 77843-3131, USA; e-mail: tferris@tamu.edu.

HUMAN FACTORS
Vol. XX, No. X, Month XXXX, pp. 1–14
DOI: 10.1177/0018720818790841

INTRODUCTION

Vehicle manufacturers are increasingly replacing physical controls (e.g., key, knob, or switch) on the center console with touchscreens. Touchscreens support a wide range of functions within a smaller spatial footprint than would be required with physical controls, accommodating context-specificity and user preferences (Irwin et al., 2011; Park & Han, 2010). One manufacturer has gone as far as to replace all of the physical controls on the center console with one 17-in. touchscreen (Brignall, 2016).

The interface transition from physical controls to touchscreens alters *nonvisual* interaction cues that help drivers better operate under visual demands. For example, drivers can feel resistance force or hear the sound of activation, and such nonvisual cues allow drivers to confirm interaction while maintaining visual awareness of the roadway. In contrast, controls displayed in touchscreens are intangible, presented on a flat surface, and thus lack the naturalistic cues that physical controls can offer. Thus, visual attention is essential for in-vehicle touchscreen interaction, while driving relies on vision, too. Driving and touchscreen interaction will therefore compete for the limited visual attentional resources (Wickens, 2002). This competition would be more challenging than that with physical controls, as touchscreen interfaces do not possess naturalistic nonvisual cues that could offload visual demands of touchscreen interaction. Such visual distraction potential would lead to more time with eyes off the road and serious concerns for driving safety (Fitch et al., 2013; Liang, Lee, & Yekhshatyan, 2012; Olson, Hanowski, Hickman, & Bocanegra, 2009).

As earlier studies generally suggest benefits of multimodal feedback in driver attention management (e.g., Lee, Roberts, Hoffman, & Angell, 2012; Ranney, Harbluk, & Noy, 2005), such benefits were also demonstrated for the design of in-vehicle touchscreens, improving drivers' confidence in touchscreen interaction (Pitts, Burnett, Williams, & Wellings, 2010), and reducing the frequency and time of glances at an in-vehicle touchscreen (Pitts, Skrypchuk, Wellings, Attridge, & Williams, 2012). However, these earlier findings were not based on actual performance measures pertaining to hazard detection, and it has been unknown whether the decreased loads in glance behavior would actually support drivers' abilities in properly and promptly responding to safety-critical objects on the roadway.

The evaluation of synthetic nonvisual feedback for in-vehicle touchscreens needs to take the context of driving into account. Lee and Spence (2008) asserted that the addition of nonvisual modalities to the feedback of touchscreen keys can reduce the subjective ratings of perceived workload of in-vehicle touchscreen tasks. They conducted a simulated driving experiment where drivers have to conduct a car avoidance task. However, the approaching cars initiated a move at a fixed interval, which made the simulated risk predictable. In order to assess drivers' abilities to maintain visual awareness of the roadway dynamics, safety-critical events that are unforeseen should be considered. Assessing such abilities can benefit from investigating performance of probe detection (Jones & Endsley, 2004; Walker, Stanton, & Young, 2008), as safety risks of drivers are first associated with visual perception of surroundings (e.g., Hyman, Boss, Wise, McKenzie, & Caggiano, 2010; Nasar & Troyer, 2013).

It is noteworthy that most previous driving studies concerning effects of perceptual modality were conducted in simulated environments (e.g., Lee & Spence, 2008; Pitts, Burnett, et al., 2012; Pitts, Skrypchuk, et al., 2012; Reimer, Mehler, & Donmez, 2014). In fact, a recent systematic review on driving studies involving touchscreens reported that only two out of 18 recent studies were conducted on the real road (Orphanides & Nam, 2017). Although simulated driving is convenient for not posing physical risks to participants, simulators would approximate multisensory experience that a real vehicle on the real road can provide to drivers. For example, the roadway and a running engine can introduce auditory noise and vibrations at various frequencies that represent naturalistic masking effects on the perception of interaction feedback or could perturb manual interaction of drivers. In this regard, more realistic testing contexts could impose contextual factors that are difficult to reproduce in simulators but may have an impact on in-vehicle interaction.

This study addressed the aforementioned issues based on (1) measuring drivers' abilities to properly and promptly respond to visually-obscured probes on the roadside and (2) employing an instrumented vehicle on a closed-course track. Drivers were distracted by a manual data entry task requiring them to use experimental dashboard interfaces. Characteristics of the interfaces varied in interface type (touchscreen or physical interface), key feedback modality (combinations of visual, auditory, and vibrotactile synthetic feedback), and key size, with the expectation that key size would not be as effective as synthetic feedback in performance or workload.

This study hypothesized that multitasking performance of drivers regarding visual probe detection and manual data entry will be better when a secondary in-vehicle task is supported by nonvisual feedback. This study also hypothesized that the effects of nonvisual feedback will reduce perceived workload. These hypotheses were based on the expectation that performance and workload would benefit from richer feedback (or more modalities) while drivers perform in-vehicle touchscreen interaction and that the best outcome would be observed with a physical interface that provides naturalistic multisensory feedback. The premise of this study concerns benefits of multimodal feedback for in-vehicle touchscreen design. The findings would be beneficial to researchers and designers whose interest is vehicle interior design in the context of driving.

METHOD

Twenty-nine drivers participated in the study (17 females, mean age = 27, range: 21–36). Their licensed driving experience was 11.2 years on average (SD = 5.8), suggesting that participants had an adequate level of experience to practice a peripheral search on the roadside (Crundall, Underwood, & Chapman, 1999; 2002). All participants demonstrated normal or corrected-to-normal

Figure 1. Overview of the experimental environment. Each roadside had 44 barrels spaced about 20 m (or 60 feet) apart.

visual acuity of at least 20/50. Informed consent was obtained from all the participants, and they received $40 for 1.5 hr of participation. This research complied with the American Psychological Association Code of Ethics and was approved by the institutional review board at Texas A&M University (TAMU IRB2014-0688F).

Driving Environment

Participants drove a 2006 Toyota Highlander on a prescribed route that was a straight and two-lane road outlined on both sides with traffic delineation barrels (TDBs) (see Figure 1). The data were collected for the time duration that participants completed driving a *lap*, which included driving the length of the straight course and driving back along the same roadway, resulting in about 1.6 kilometers (or 5,200 feet) with lane changes (turning around was excluded). Participants took approximately 3 min to complete one lap, or the two straight routes.

A high-definition camera hung on the ceiling of the interior recorded the video, as illustrated in Figure 2. The resolution was 1920 × 1080, and the frame rate turned out to be 23 frames per s on average. The camera also recorded audio to confirm verbal responses from participants.

Experimental Tasks

Participants drove the vehicle concurrently performing a set of experimental tasks consisting of (1) an input task using keypads configured with various interface characteristics, and (2) a detection task to verbally acknowledge and promptly respond to visual probes obscured by the barrels placed along the roadside. Participants were instructed that all tasks were of equivalent priority.

Input task. This task required entering predefined input sequences with each of the keypads shown in Figure 3: (1) a physical keypad (referred to as “Physical”), (2) a touchscreen-based keypad with small keys (“TouchSmall”) that were arranged in the same dimensions as Physical, and (3) a touchscreen-based keypad that was enlarged and proportionately spaced (“TouchLarge”). Physical was affixed to a hardboard in a location that mapped to the key locations of TouchSmall. The numbers printed on Physical were not used, instead participants were instructed to consider it as a 3 × 3 key array with “enter” (the long key at the bottom) and “clear” (the one next to “enter”) keys in the bottom row.

Table 1 presents the size specification of the three keypads. The key size and spacing of TouchSmall were intended to be identical to those of the physical keys, while TouchLarge were designed to be arguably optimal in size and spacing for discrete selection tasks (Wickens, Gordon, Liu, & Lee, 1998).

Figure 4 shows an illustration of the input sequence instructions. With a smartphone screen mounted above the instrument cluster, the instructions visually

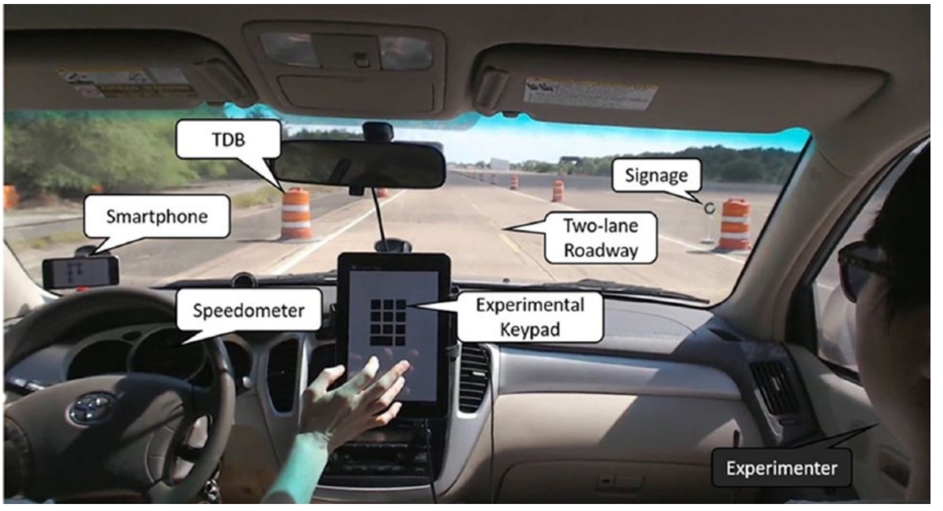

Figure 2. View of the experimental setting captured via HD video camera. Participants visually attended to at least six objects including the TDBs, the smartphone, the speedometer, the two-lane roadway, one of the three experimental keypads, and the signs.

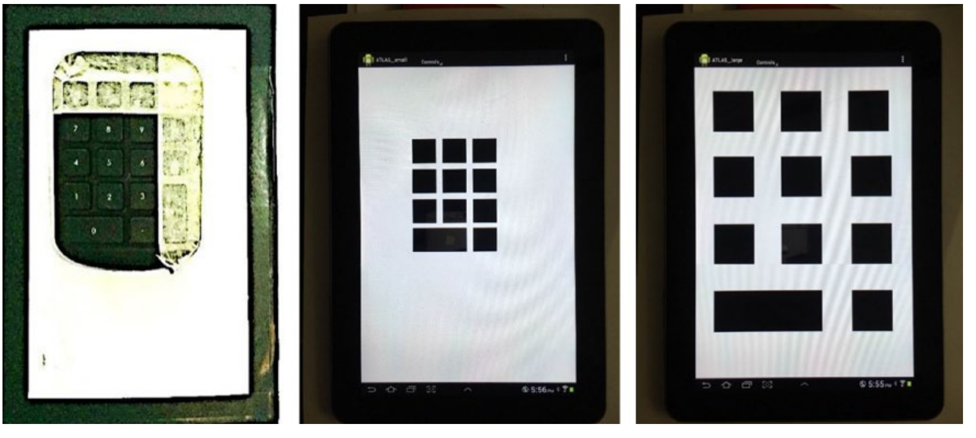

Figure 3. Experimental interfaces—Physical (left), TouchSmall (middle), and TouchLarge (right). Physical and TouchSmall were regarded as identical in size, and the unnecessary keys of Physical were taped and masked. TouchSmall and TouchLarge were different in size and had different configurations of synthetic feedback.

presented as 5-key input sequences that included vertically- or horizontally-adjacent keys (without repeat). The beginning of sequences was represented with a diamond shape; an arrow indicated the end. To register one input sequence, participants pressed the corresponding five keys in the depicted order, followed by the “enter” key. They were also instructed to press the “clear” key when they thought they made an errant entry, but none of the keypads displayed which key was inputted. Each 5-key sequence was designed to be completed within 2 s, based on the recent National Highway Traffic Safety Administration (NHTSA) guideline concerning visual-manual interaction while driving (NHTSA, 2014). Upon the completion of one sequence, participants touched the smartphone screen, at the earliest convenience, to advance to the next input sequence. They

TABLE 1: Specifications of Key Size and Spacing (mm)

Keypad type	Input and Clear Keys		Enter Key		Spacing Between Keys
	Width	Height	Width	Height	
Physical and TouchSmall	15	15	34	15	4.5
TouchLarge	25	25	67	25	17

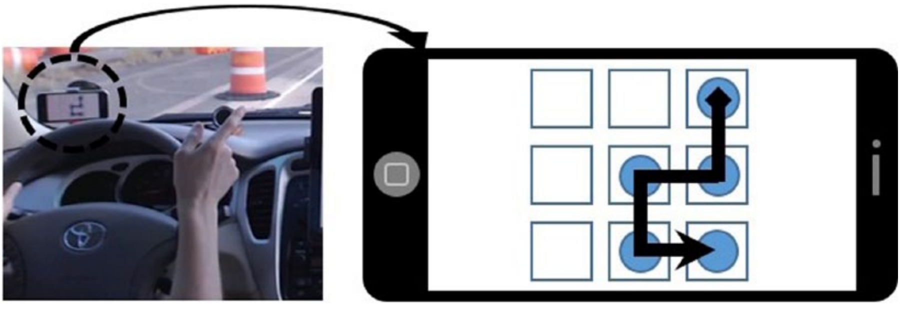

Figure 4. An illustration of a graphical instruction for the input task. With the smartphone located on the left side of the windshield, participants had to frequently reorient visual attention to interact with the experimental keypads on the center console. A unique set of fifteen instructions were associated with each of the six laps, and the set was repeated for the duration of the lap.

continued to renew the sequences and to input them until completing a lap, encouraged to register as many correct sequences as possible.

Table 2 describes six interface characteristics (the independent variable) and three types of synthetic feedback applied to touchscreen conditions. None of the participants reported a problem with perceiving any type of synthetic feedback.

Detection task. This task required participants to verbally acknowledge signs placed between the barrels and then to immediately perform a lane change only if instructed by the acknowledged sign. Figure 5 shows the signage which consisted of incomplete circles printed on transparent plastic boards. The plastic boards were affixed to a delineation post, and it was possible to rotate them about a pivot point to change orientations of the sign. The “C” orientation (the left of Figure 5) signified a need to Change lanes immediately following a verbal response, whereas the “U” orientation (the right of Figure 5) required a verbal response only (“lane should be Unchanged”). Participants were allowed to use their own words for verbal responses, but encouraged to simply say “C”/“Change” or “U”/“Unchanged.”

Eight signs were placed on each roadside. Sight lines to the signs were controlled in a way that signs were only visible when approaching the obscuring barrel. Each sign was placed 1.5 m away from the nearest barrel at a consistent offset so that the sign can reveal itself and allow participants a fairly consistent amount of time to detect the sign. Figure 6 shows screenshots from the recorded video to illustrate the unveiling of one of the signs.

Experimenters in a separate, trailing vehicle pseudo-randomly reconfigured the orientation of the signs and the number of barrels between signs before starting each lap, under the only condition that the number of barrels between signs were in the range between three and seven (approximately 300 to 700 feet).

Procedure

The training sessions were conducted on the same course track as the one used for data collection. As mentioned earlier, each lap consisted

TABLE 2: Interface Characteristics and Types of Feedback Configuration

Category	Type	Description
Interface Characteristics	TouchSmall-V	TouchSmall with visual feedback only
	TouchSmall-VA	TouchSmall with visual and auditory feedback
	TouchSmall-VT	TouchSmall with visual and vibrotactile feedback
	TouchSmall-VAT	TouchSmall with visual, auditory, and vibrotactile feedback
	TouchLarge-VAT	TouchLarge with visual, auditory, and vibrotactile feedback
	Physical	The physical keypad
Synthetic Feedback	V (Visual)	The key color blinks from black to yellow (RGB: #FFFF00) when touched
	A (Auditory)	The keystroke sound of the tablet's default setting is played
	T (vibroTactile)	The tablet vibrates for 0.5 s

Note. The intensity of auditory and vibrotactile feedback was set to be the tablet's maximum levels.

Figure 5. Experimental signage (Left: "C" orientation, Right: "U" orientation). The "C" orientation signified a need to Change lanes following a verbal response, whereas the "U" orientation required a verbal response only ("lane should be Unchanged"). The post heights were adjusted to be as tall as the TDBs so the post and the affixed sign can be obscured in distance. The signage was made in two sizes: 10-in. and 6-in. diameter. Eight signs on each roadside included four pairs of 10-in. and 6-in. signs, alternating the large and small sizes.

of two straight routes; participants drove down and back along the straight road. The training involved four laps. During the first lap, they conducted the input task in the first half of the lap and the detection task in the second half. For the remaining three laps, participants conducted both tasks concurrently with each of the six interface settings from Table 2. Each interface setting was mapped to each of six straight routes (three laps in total).

In the main experiment, participants concurrently performed the experimental tasks. They completed six laps to test each of the six interface settings (see Table 2). The order in which

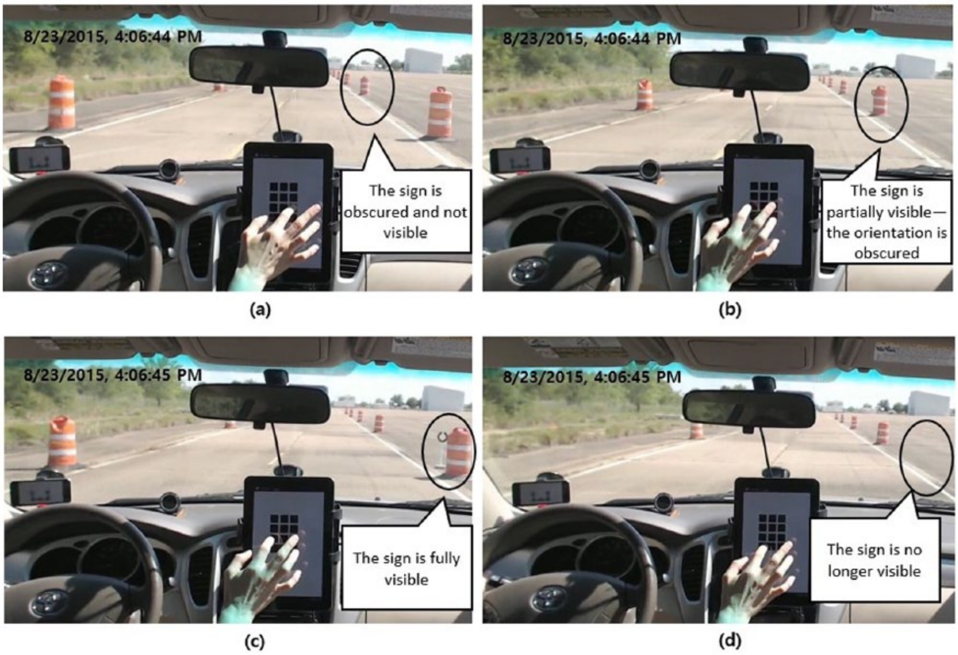

Figure 6. An illustration of signage detection in the detection task. Although the vehicle was approaching the TDB (a), the participant could not acknowledge the sign until sight lines allowed the sign to be unveiled (b). In a brief moment, the orientation became visible (c). When traveling at 25 mph, each sign became no longer visible in less than 2 s (d). Participants were instructed to monitor the right side of the roadway for the detection task.

the six settings were used was partially counterbalanced in order for each condition to appear once in each position of the order, like a standard Latin Square Design. Regarding the vehicle speed, participants were instructed to maintain 40 kph (or 25 mph) throughout each lap. Before entering the course, 60 m were given to accelerate to 40 kph. Participants were told that no notification will be given for speed violation during the main experiment.

Measures

For the input task, *input efficiency* was defined as the quantity of all correctly entered sequences. *Input accuracy* (%) was defined as the percentage of instruction sequences that were entered correctly. Cleared inputs (pressing "clear" to start over the entry before pressing "enter") were not included. *Input time* (ms) was the mean average of time from pressing the first key of a sequence until pressing enter. Only correctly entered sequences were considered in analysis.

For the detection task, the video and audio recordings were analyzed; both verbal and manual responses needed to be correct to be registered as "correct response." *Detection rate* (%) was the percentage of the correct responses over all the signs.

The assessment of temporal performance of the detection task was challenging, as the recorded video did not display elapsed time. Moreover, the time at which a sign became visible to participants varied; the sight lines to the sign differed depending on the driving lane in which the vehicle was traveling at the moment. Hence, a unique measure was needed to assess temporal performance, and this measure needed to be independent of driving lane positions. *Response promptness* (%) represents the time at which a participant correctly acknowledged a given sign, evidenced by the view available to the ceiling-mounted video camera (see Figure 2). Figure 7 depicts the concept of response promptness; "A" represents the time elapsed

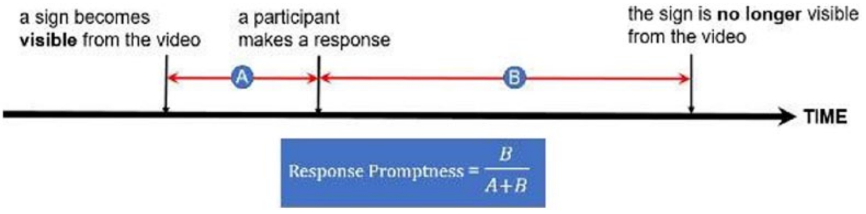

Figure 7. The concept of response promptness.

between the emergence of a given sign and the initiation of the verbal response, and "B" represents the time measured from the response initiation until the acknowledged sign disappeared from the recorded view. The measure of time was assessed as a count of video frames. *Response promptness* therefore was defined as the percentage ratio of video frames corresponding to B and the sum of A and B (see Figure 7). For example, a response promptness score of 40% could mean that 20 video frames elapsed between verbal acknowledgement of a particular sign and the final frame of sign visibility, given that the total duration of sign visibility was 50 frames. For missing or incorrect responses, the measure was coded as 0%.

For *perceived workload* (%), participants recorded the subjective ratings of perceived workload based on NASA-Task Load Indices (Hart & Staveland, 1988), upon the completion of each lap. On a 9-point scale, participants rated the six dimensions consisting of mental, physical, and temporal demand, performance, effort, and frustration. The collected data include the mean percentage of the 9-point ratings of the six dimensions for each participant in each condition.

Analysis

With the interface characteristics from Table 2 as the within-subject variable, each measure was compared in repeated measures analyses of variance (ANOVAs). When significance was found, post-hoc pairwise comparison was conducted with the Bonferroni adjustment. One measure, *input efficiency*, was log-transformed to avoid violating the normality assumption of ANOVA. In cases where the data exhibited substantial skewness or kurtosis, the power of an F-test was assumed to be not valid (West, Finch, & Curran, 1995), and the nonparametric Kruskal-Wallis test was used in place of ANOVA. All statistical analyses used a significance level of alpha = .05.

RESULTS

Figure 8 shows the mean values of performance measures in the input task; the letters represent significantly different groups. For *input accuracy*, skewness was substantial (−2.075) (West et al., 1995), and thus the Kruskal-Wallis test was used. The *input efficiency* data were log-transformed to satisfy the normality assumption (verified by the Shapiro-Wilk test: $W = .991, p = .349$), and the variance homogeneity was verified with Levene's test, $F(5, 168) = .508, p = .770$.

Interface characteristics significantly affected input accuracy (top left of Figure 8), χ^2 $(5, n = 29) = 17.401, p = .004, W = .77$. The post-hoc test with Bonferroni's adjustment revealed that the input accuracy of TouchSmall-V (M = 90.5%, $p = .006$, Cohen's $d = 0.846$) and TouchSmall-VT (90%, $p = .003$, $d = 1.005$) were lower than that of TouchLarge-VAT, which exhibited the best mean at 97% accuracy. The log-transformed input efficiency data were significantly affected by interface characteristics, $F(5, 168) = 3.432, p = .001$, partial $\eta^2 = .095$. The post-hoc test suggested that the input efficiency of TouchSmall-V (M=23) was significantly lower than that of Physical (29, $p < .001$, $d = 0.789$). The effect of interface characteristics on *input time* did not reach significance, $F(5, 168) = 2.0375, p = .076$.

Figure 9 shows the mean values of performance measures in the detection task. Some of the data were lost due to technical problems with video recording. The remaining data include 27 participants for the model of detection rate and 25 for the model of response promptness. Although the Shapiro-Wilk test indicated violation of the

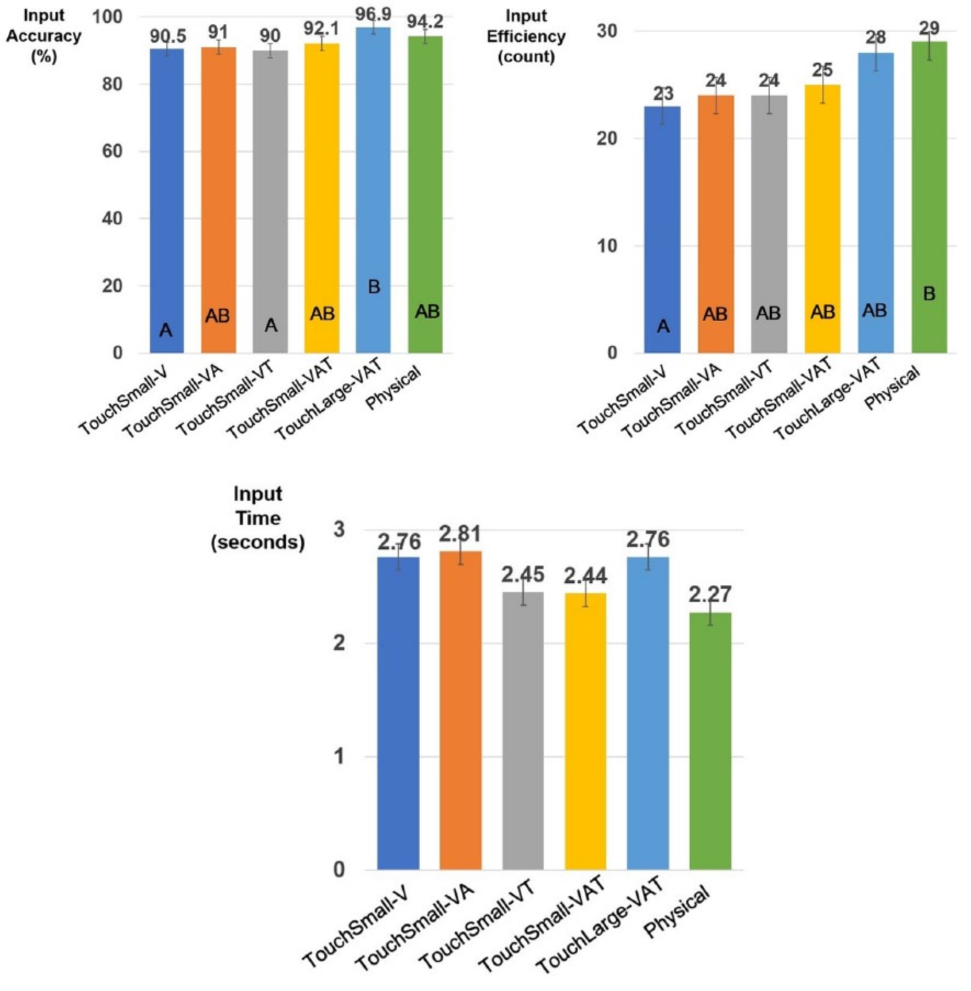

Figure 8. Performance measures of the input task (n = 29): (a) input accuracy (%) (top-left), (b) input efficiency (count) (top-right), and (c) input time (seconds) (bottom). *Input efficiency* in this figure was not log-transformed. Error bars represent the standard errors, and letters on the mean values are used to label significantly different groups. For example, A is significantly different from B but not different from AB, as AB includes A.

normality assumption for detection rate, the power of the F-test was assumed to be valid, with moderate skewness and kurtosis of −1.384 and 2.657, respectively (Harwell, Rubinstein, Hayes, & Olds, 1992; Lix, Keselman, & Keselman, 1996; West et al., 1995). Levene's test verified variance homogeneity, $F(5, 156) = 1.495$, $p = .195$. For *response promptness*, normality and variance homogeneity of the data were verified by the Shapiro-Wilk test, $W = .986$, $p = .141$, and by Levene's test, $F(5, 144) = 0.389$, $p = .856$, respectively.

Interface characteristics affected detection rate, $F(5, 156) = 6.705$, $p < .001$, partial $\eta^2 = .177$, and the post-hoc test with Bonferroni's adjustment reported that detection rate was significantly lower in TouchSmall-V (M=84.8%) than in four other conditions: TouchSmall-VT (94.9%, $p < .001$, $d = 1.170$), TouchSmall-VAT (95.3%, $p < .001$, $d = 1.273$), TouchLarge-VAT (93.7%, $p < .001$, $d = 0.936$), and Physical (94.6%, $p < .001$, $d = 1.188$). Response promptness was also affected by interface characteristics, $F(5, 144) = 5.220$, $p < .001$, partial $\eta^2 = .154$, and mirrored the post-hoc results on detection rate in that the post-hoc test showed significantly poorer performance with TouchSmall-V (the

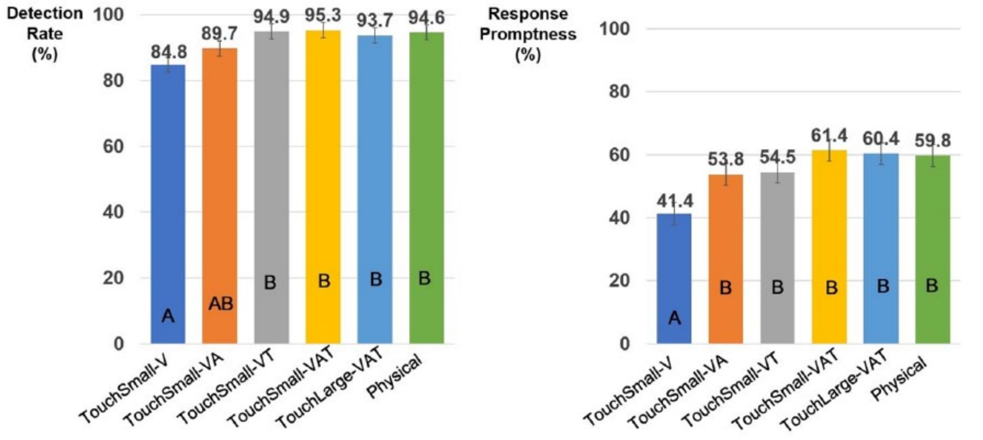

Figure 9. Performance measures of the acknowledgment task (left: detection rate [%, n = 27], right: response promptness [%, n = 25]). Error bars represent the standard errors, and letters on the mean values are used to label significantly different groups.

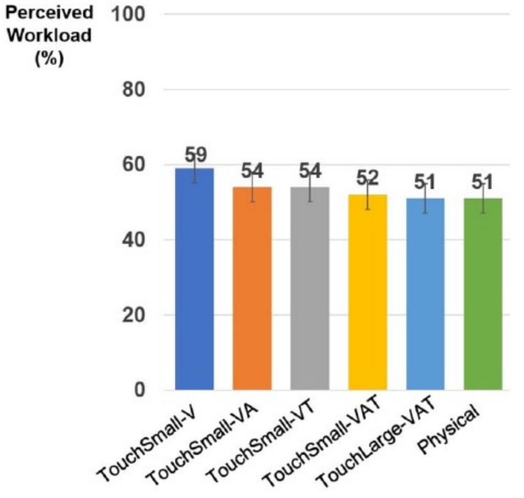

Figure 10. Perceived workload (%, n = 29). Error bars represent standard error. No significant effect was found.

lowest M = 41.4%) compared with TouchSmall-VA (53.8%, p = .005, d = 1.067), TouchSmall-VT (54.5%, p = .006, d = 1.053), TouchSmall-VAT (61.4%, p < .001, d = 1.860), to TouchLarge-VAT (60.4%, p < .001, d = 1.577), and to Physical (59.8%, p < .001, d = 1.590).

Figure 10 shows the collected data on *perceived workload*. Not enough evidence indicated a significant effect.

DISCUSSION

With automakers incorporating various non-driving functions into vehicles, touchscreens offer opportunities to provide advanced control functionality within a spatially constrained environment. However, the engagement of touchscreen interaction while driving develops multitasking situations that can cause a competition for limited attentional resources (Wickens, 2002), especially with regard to visual resource demands. The design of in-vehicle touchscreens should consider distraction potentials regarding the costs imposed on visual attention management while driving, and design decisions should be informed based on investigations about how interface characteristics can impact those costs.

Previous research demonstrated that driver distraction can be associated with design factors of in-vehicle touchscreens including but not limited to interaction feedback (Pitts et al., 2010), information search methods (Kujala, 2013), and menu configuration styles (Kujala & Salvucci, 2015). The findings of the current research build on this body of work to show how interface characteristics, including interface type, key feedback modality, and key size, can affect performance of visual detection and touchscreen interaction. The two concurrent tasks that participants performed created the competition for visual processing resources. This task situation required drivers to regulate the level of risk by switching between the input task and the visual detection task (Kujala & Salvucci, 2015), and performance was affected by the availability of nonvisual feedback from the input interface. When participants were not supported by nonvisual feedback that would have been useful to confirm the activation of input keys, visual attention on the roadway needed to be reoriented to the secondary task interface more frequently (Pitts et al., 2012). The collected data show that performance of visual detection was contingent upon the feedback configuration.

First of all, synthetic nonvisual feedback from the touchscreen significantly supported visual awareness of the road. The detection task required participants to allocate visual attention to the road as often as possible, and the worst performance in detection rate and in response promptness was observed when the in-vehicle touchscreen interaction was relying on unimodal visual feedback. Most of the touchscreen conditions with nonvisual feedback showed higher detection rate and response promptness, compared with the touchscreen condition with unimodal visual feedback. Pitts et al. (2012) asserted that drivers can glance at in-vehicle touchscreens more frequently and longer in time when the touchscreen interaction is not supported by "haptic" feedback. However, it has been unknown how the changed glance behavior associated with the quality of touchscreen feedback would undermine performance of drivers. The current findings illustrate how the absence of nonvisual modalities of in-vehicle touchscreen interaction can incur additional costs on visual awareness of the road, evidenced by actual performance measures concerning accuracy and time. In the same regard, the findings also show the extent to which synthetic nonvisual feedback can reduce those costs. Multiple resource theory (Wickens, 2008) explains these findings, as the addition of nonvisual cues to touchscreen interaction can offload visual demands, and the task interferences due to the limited visual attention can be alleviated.

It is noteworthy that the interface transition from physical keys to touchscreen ones resulted in a significant difference when the touchscreen was not supported by synthetic nonvisual feedback; effects were significant for detection rate, response promptness, and input efficiency. However, the physical keypad, involving the richest form of feedback, was not significantly different from the touchscreen conditions in performance of the detection task, when any nonvisual modalities were available to the touchscreen conditions. These findings suggest that the distraction potential involved in the interface transition can be addressed by a simple method, to a degree; the addition of auditory notification or vibration can be helpful for reducing visual demands of touchscreen interaction. Moreover, it would be more cost-effective than employing proprietary forms of "haptic" feedback (e.g., Lee & Spence, 2008; Pitts et al., 2012).

Regarding the input task, the results indicate that the interface characteristics significantly affected input accuracy and input efficiency. However, not enough evidence was found to claim that the addition of nonvisual feedback can effectively support touchscreen interaction. Given the small size of touchscreen keys (the TouchSmall conditions), pairwise comparisons did not indicate a significant difference in performance measures of the input task. This was in part consistent with an earlier study reporting no significant effect in input time (Lee & Spence, 2008).

The effects of key size were tested, based on the comparison between TouchSmall-VAT and TouchLarge-VAT. This pairwise comparison did not show significance in any of the measures. Nevertheless, one should not conclude that the effects of key size do not matter in in-vehicle touchscreens. The current study employed key sizes of 15 mm and 25mm, and both might have been too large to expect a difference. In a similar regard, an earlier study (Kim, Kwon, Heo, Lee, & Chung, 2014) reported that key sizes larger than 17.5 mm did not show significance in input error, input time, user preference, and driving performance.

Despite an initial assumption that synthetic feedback configuration would affect perceived workload, the results do not provide evidence to support it. On the contrary, Lee and Spence (2008) found a significant effect in the same measure. The earlier study used the identical assessment technique as the one used in the current study, suggesting that perceived workload was significantly higher in the unimodal visual feedback condition compared with the other conditions. However, the current study did not find such an effect. Regarding this discrepancy in the statistical results, driving situations that the two studies involved might have an explanation. Both studies required participants to change driving lanes, but instructions for doing so were presented in a different way. Lee and Spence (2008) involved a car avoidance task in which drivers had to change lanes to avoid *simulated crash risks*, whereas the present study involved *benign signs*. In this regard, future studies may first examine whether such effects on perceived workload actually depend on driving situations, with a two-way ANOVA design based on the combinations of feedback modalities and driving situations. Such a design can offer opportunities to investigate an interaction effect associated with feedback modalities.

Overall, the findings of this study suggest that the richness of synthetic feedback of touchscreens can affect driver visual attention management and performance, which could contribute to safety guidelines for in-vehicle interface design. For instance, the NHTSA (2014) published a guideline involving safety concerns for driver distraction due to in-vehicle electronic devices, including touchscreens. Calling for automakers to limit the amount of time associated with in-vehicle touchscreen interactions, the guideline lacks suggestions relating to distraction mitigation methods such as multisensory feedback. The current findings emphasize that auditory or tactile feedback, or both, can support drivers for the management of visual attention and affect performance. Touchscreens offer designers the flexibility of interaction design and the efficient use of the limited space, yet designers need to be well informed about how touchscreen interface characteristics impact performance of drivers in safety. Otherwise, a common pattern can be realized, introducing driver-support features that inadvertently increase safety risks by becoming a source of distraction (Petzoldt, Bellem, & Krems, 2014).

The current study involves limitations. First, the number of participants was relatively small, although the effect sizes were considerably large for significant effects to be found. Second, the design of the experiment was not full factorial; the large touchscreen keys involved only one feedback condition whereas the small-key conditions were compared based on different feedback modalities. Therefore, in some cases, the confounding effects originating in the combination of key feedback modality and key size made it difficult to assess individual interface characteristics exclusively. For example, significant differences in input accuracy were found when comparing TouchLarge-VAT to TouchSmall-V and to TouchSmall-VT, but it is difficult to determine whether the differences are attributable to key size or key feedback modality. Third, benefits of naturalistic feedback from physical keys were not fully appreciated in this study. Such naturalistic feedback would have provided drivers with haptic cues (e.g., shape and texture) to locate keys without vision. However, participants in this study were not explicitly trained or instructed to take advantage of such cues before pressing keys, as the premise of this study was about effects of key feedback upon the activation of keys, not about the cues for searching keys. As a result, not all participants would have exploited naturalistic "pre-input" cues from physical keys. It is also noticeable that the current findings do not involve performance of vehicle operation; lateral positions maintenance or speed maintenance would have been other indicators of distraction effects. Lastly, the present study did not involve a treatment condition with no secondary task, which would have been useful to validate the results in terms of perceived workload.

CONCLUSION

This study analyzed multitasking performance and safety aspects with respect to interface characteristics for in-vehicle touchscreens and a physical keypad mounted on the center console. Participants conducted a lane-change task on a closed-course track, according to the signs placed on the roadside. They were concurrently tasked with the experimental keypads on the center console that were differed by interface characteristics including interface type, key feedback modality, and key size. The findings suggest that the addition of nonvisual feedback to touchscreen keypads can better support performance of visual search in terms of detection rate and response promptness. The findings also illustrate that such nonvisual feedback can alleviate the performance degradation observed in the interface transition from physical keypad to touchscreens, informing design decisions for vehicle manufacturers. Further research should involve experimental driving conditions that more readily encourage naturalistic driving behavior.

ACKNOWLEDGMENTS

This research project was supported by the Center for Advancing Transportation Leadership and Safety (ATLAS Center). The ATLAS Center is supported by a grant from the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Transportation, University Transportation Centers Program (DTRT13-G-UTC54). The ATLAS Center is a collaboration between the University of Michigan Transportation Research Institute and the Texas A&M Transportation Institute. Authors would like to extend gratitude to ATLAS representatives, Texas A&M Transportation Institute researchers, and assistant staffs, especially including Jeffrey Miles, Christine Yager, Ko-Ching (Katrina) Wu, and Sungeun Cho for their roles in study design, project management, and technical support activities.

KEY POINTS

- Sensory feedback issued from in-vehicle dashboard interfaces can affect drivers in interaction quality and attention management.
- Driver interaction with a touchscreen-based dashboard can be more accurate and efficient when the touchscreen provides both visual and synthetic nonvisual feedback.
- Performing a secondary in-vehicle interaction task can lead to performance degradation of visual search on the roadway, and performance is contingent upon the feedback condition of the task interface.
- Physical control elements can outperform touchscreens in supporting multitasking performance, though the difference in performance is less consequential when touchscreens include synthetic multisensory feedback.

ORCID ID

Youngbo Suh https://orcid.org/0000-0001-9777-945X

REFERENCES

Brignall, M. (2016, May 2). Is the Tesla 3 the electric car that will change everything? *The Guardian*. Retrieved from http://www.theguardian.com/money/2016/may/07/tesla-3-electric-car-affordable-worth-price.

Crundall, D., Underwood, G., & Chapman, P. (1999). Driving experience and the functional field of view. *Perception*, *28*(9), 1075–1087.

Crundall, D., Underwood, G., & Chapman, P. (2002). Attending to the peripheral world while driving. *Applied Cognitive Psychology*, *16*(4), 459–475.

Fitch, G. M., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., Olson, R. L., Perez, M. A., Hanowski, R. J., Hankey, J. M., & Dingus, T. A. (2013). The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk. DOT HS 811757, Washington, DC: U.S. Department of Transportation, National Highway Traffic Safety Administration.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, *52*, 139–183.

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo results in methodological research: The one-and two-factor fixed effects ANOVA cases. *Journal of Educational and Behavioral Statistics*, *17*(4), 315–339.

Hyman, I. E., Boss, S. M., Wise, B. M., McKenzie, K. E., & Caggiano, J. M. (2010). Did you see the unicycling clown? Inattentional blindness while walking and talking on a cell phone. *Applied Cognitive Psychology*, *24*(5), 597–607.

Irwin, C. B., Yen, T. Y., Meyer, R. H., Vanderheiden, G. C., Kelso, D. P., & Sesto, M. E. (2011). Use of force plate instrumentation to assess kinetic variables during touch screen use. *Universal Access in the Information Society*, *10*, 453–460.

Jones, D. G., & Endsley, M. R. (2004). Use of real-time probes for measuring situation awareness. *The International Journal of Aviation Psychology*, *14*(4), 343–367.

Kim, H., Kwon, S., Heo, J., Lee, H., & Chung, M. K. (2014). The effect of touch-key size on the usability of In-Vehicle Information Systems and driving safety during simulated driving. *Applied Ergonomics*, *45*(3), 379–388.

Kujala, T. (2013). Browsing the information highway while driving: Three in-vehicle touch screen scrolling methods and driver distraction. *Personal and Ubiquitous Computing*, *17*(5), 815-823.

Kujala, T., & Salvucci, D. D. (2015). Modeling visual sampling on in-car displays: The challenge of predicting safety-critical lapses of control. *International Journal of Human-Computer Studies*, *79*, 66–78.

Lee, J. D., Roberts, S. C., Hoffman, J. D., & Angell, L. S. (2012). Scrolling and driving: How an MP3 player and its aftermarket controller affect driving performance and visual behavior. *Human Factors*, *54*(2), 250–263.

Lee, J.-H., & Spence, C. (2008). Assessing the benefits of multimodal feedback on dual-task performance under demanding conditions. *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*, *1*, 185–192.

Liang, Y., Lee, J. D., & Yekhshatyan, L. (2012). How dangerous is looking away from the road? Algorithms predict crash risk from glance patterns in naturalistic driving. *Human Factors:*

The Journal of the Human Factors and Ergonomics Society, *54*(6), 1104–1116.

Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research*, *66*(4), 579–619.

Nasar, J. L., & Troyer, D. (2013). Pedestrian injuries due to mobile phone use in public places. *Accident Analysis and Prevention*, *57*, 91–95.

National Highway Traffic Safety Administration. (2014). Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices, *Federal Register*, *78*(81), 55530–55534.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). Driver distraction in commercial vehicle operations. Report No. FMCSA-RRR-09-042. Washington, DC: U.S. Department of Transportation, Federal Motor Carrier Safety Administration.

Orphanides, A. K., & Nam, C. S. (2017). Touchscreen interfaces in context: A systematic review of research into touchscreens across settings, populations, and implementations. *Applied Ergonomics*, *61*, 116–143.

Park, Y. S., & Han, S. H. (2010). Touch key design for one-handed thumb interaction with a mobile phone: Effects of touch key size and touch key location. *International Journal of Industrial Ergonomics*, *40*, 68–76.

Petzoldt, T., Bellem, H., & Krems, J. F. (2014). The critical tracking task: A potentially useful method to assess driver distraction? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *56*(4), 789–808.

Pitts, M. J., Burnett, G. E., Williams, M. A., & Wellings, T. (2010, November). Does haptic feedback change the way we view touchscreens in cars? *12th International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, 38. Beijing

Pitts, M. J., Burnett, G., Skrypchuk, L., Wellings, T., Attridge, A., & Williams, M. A. (2012). Visual–haptic feedback interaction in automotive touchscreens. *Displays*, *33*(1), 7–16.

Pitts, M. J., Skrypchuk, L., Wellings, T., Attridge, A., & Williams, M. A. (2012). Evaluating user response to in-vehicle haptic feedback touchscreens using the lane change test. *Advances in Human-Computer Interaction 2012*, 1–13.

Ranney, T. A., Harbluk, J. L., & Noy, Y. I. (2005). Effects of voice technology on test track driving performance: Implications for driver distraction. *Human Factors*, *47*(2), 439–454.

Reimer, B., Mehler, B., & Donmez, B. (2014). A study of young adults examining phone dialing while driving using a touchscreen vs. a key style flip-phone. *Transportation Research Part F: Traffic Psychology and Behaviour*, *23*, 57–68.

Walker, G. H., Stanton, N. A., & Young, M. S. (2008). Feedback and driver situation awareness (SA): A comparison of SA measures and contexts. *Transportation Research Part F: Traffic Psychology and Behaviour*, *11*(4), 282–299.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 56–75). Thousand Oaks, CA: SAGE.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2): 159–177.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 449–455.

Wickens, C. D., Gordon, S. E., Liu, Y., & Lee, J. (1998). *An introduction to human factors engineering*. Upper Saddle River, NJ: Prentice Hall, Inc.

Youngbo Suh is an associate research fellow of Korea Institute for Defense Analyses in Seoul, South Korea. He earned his PhD in Industrial and Systems Engineering from Texas A&M University in 2018. He received an MS in 2007 and a BS in 2005, both from Industrial Engineering at Seoul National University, Seoul, South Korea.

Thomas K. Ferris is an associate professor of Industrial and Systems Engineering at Texas A&M University, where he is the director of the Human Factors & Cognitive Systems Laboratory. He earned his PhD in Industrial and Operations Engineering from the University of Michigan in 2010.

Date received: October 13, 2017
Date accepted: July 2, 2018