International Journal of Mobile Human Computer Interaction
Volume 9 • Issue 2 • April-June 2017

Smartwatches vs. Smartphones:
Notification Engagement while Driving

Wayne C.W. Giang, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada
Huei-Yen Winnie Chen, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada
Birsen Donmez, Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada

ABSTRACT

This work seeks to understand whether the unique features of a smartwatch, compared to a smartphone, mitigate or exacerbate driver distraction due to notifications, and to provide insights about drivers' perceptions of the risks associated with using smartwatches while driving. As smartwatches are gaining popularity among consumers, there is a need to understand how smartwatch use may influence driving performance. Previous driving research has examined voice calling on smartwatches, but not interactions with notifications, a key marketed feature. Engaging with notifications (e.g., reading and texting) on a handheld device is a known distraction associated with increased crash risks. Two driving simulator studies compared smartwatch to smartphone notifications. Experiment I asked participants to read aloud brief text notifications and Experiment II had participants manually select a response to arithmetic questions presented as notifications. Both experiments investigated the resulting glances to and physical interactions with the devices, as well as self-reported risk perception. Experiment II also investigated driving performance and self-reported knowledge/expectation about legislation surrounding the use of smart devices while driving. Experiment I found that participants were faster to visually engage with the notification on the smartwatch than the smartphone, took longer to finish reading aloud the notifications, and exhibited more glances longer than 1.6 s. Experiment II found that participants took longer to reply to notifications and had longer overall glance durations on the smartwatch than the smartphone, along with longer brake reaction times to lead vehicle braking events. Compared to the no device baseline, both devices increased lane position variability and resulted in higher self-reported perceived risk. Experiment II participants also considered that smartwatch use while driving deserves penalties equal to or less than smartphone use. The findings suggest that smartwatches may have road safety consequences. Given the common view among participants to associate smartwatch use with equal or less traffic penalties than smartphone use, there may be a disconnect between drivers' actual performance and their perceptions about smartwatch use while driving.

KEYWORDS

Distracted Driving Laws, Driver Distraction, Risk Perception, Smartphones, Smartwatches

DOI: 10.4018/IJMHCI.2017040103

1. INTRODUCTION

Smartphones, one of the most popular consumer electronic devices today, go beyond traditional phone functions such as texting and calling, and provide continuous access to social media, current news, and many other information sources. The extensive amount of time that many spend on their smartphones has led to habits and even addictions formed towards smartphone use (e.g., Bayer & Campbell, 2012; Oulasvirta, Rattenbury, Ma, & Raita, 2012; Walsh, White, & Young, 2010); in one study, participants were found to check their smartphones on average 34 times a day (Oulasvirta et al., 2012). Such dependency on smartphones is problematic when these devices are brought into the vehicle. At any given time, an estimated 1.7% of U.S. drivers are manipulating their handheld devices on the road (Pickrell, 2015), and naturalistic driving studies reveal that visual-manual interactions with these devices in particular lead to a major increase in crash risks (Fitch et al., 2013; Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006).

New forms of smart, connected devices that are finding their ways into the vehicle can also create use patterns that may lead to safety decrements. A notable trend over the past few years is the marketing of wearable technologies such as smartwatches and Google Glass as consumer products. As devices that are worn or attached to the body, wearables are highly portable and often make use of multiple input modalities (e.g., touch, voice, or gesture), thus making their functions even more accessible to drivers on the road. Smartwatches, in particular, are rapidly gaining popularity among consumers: 19 million units are estimated to ship in 2015 and 103 million units in 2019 (Danova, 2015). Worn on the user's wrist like a traditional watch, smartwatches are connected to the user's smartphone via Bluetooth for accessing notifications about incoming emails, text messages, etc. Some smartwatches also provide additional functions such as initiating or taking calls. Given the expanding smartwatch user base and the variety of functions smartwatches provide, there is a need for understanding how drivers interact with smartwatches and what potential consequences smartwatches may have on road safety.

To the best of our knowledge, in addition to the two experiments we report in this paper, only one other study so far has examined smartwatch use while driving. Samost et al. (2015) used a simulated car following task, combined with a remote detection response task, to compare smartwatch and smartphone use for initiating phone calls. No driving performance differences were found for voice calling (auditory-visual), a somewhat expected result given that the voice input methods on the two devices were similar. Compared to voice calling on either device, visual-manual calling on the smartphone led to worse driving performance in terms of lane deviation and major steering wheel reversals. While Samost et al. (2015) focused on the voice calling aspect of a smartwatch, the most popular function of smartwatches is the notification system. Positioned on their wrist, smartwatches allow users to filter their incoming notifications without having to reach for their smartphone (Marks, 2013).

Whether from a smartphone or a smartwatch, notifications carry the risk of diverting attention away from the main driving task. Distraction due to notifications may have two components. On one hand, drivers may intentionally engage with notifications (e.g., texting a response to an incoming message), a visual-manual task that is known to increase crash risk considerably when performed on a handheld phone (Fitch et al., 2013). On the other hand, with or without intending to engage with notifications, drivers may be involuntarily affected by the auditory, visual, and/or tactile alerts that accompany notifications (Marulanda, Chen, & Donmez, 2015; Regan, Hallett, & Gordon, 2011). The sudden onset of peripheral visual stimuli (without message content), for example, has been shown to attract visual attention and degrade accelerator release times in response to lead vehicle braking events (Hoekstra-Atwood, 2015; Hoekstra-Atwood, Chen, Giang, & Donmez, 2014).

The unique characteristics of smartwatches may particularly accentuate the involuntary distraction associated with receiving notifications. The proximity of the smartwatch to the user (i.e., on the wrist rather than elsewhere in the vehicle) and the vibrotactile feedback that smartwatches provide directly onto the body may elevate the perceived intensity of notification alerts. This increase in perceived intensity coupled with the ease of access to notification content (e.g., a simple flick of the wrist can turn the backlight on) can increase the likelihood of an involuntary, automatic response, and hence driver distraction.

As mentioned previously, no research has examined driver engagement with notifications on smartwatches compared to smartphones. To provide a preliminary assessment of the distinction between the two devices in terms of their distraction potential, we present two driving simulator studies comparing notification engagement on smartwatches to smartphones. The first study was exploratory with six participants. We constructed a primarily visual task, where participants were asked to read aloud notifications (each containing a simple address) while they drove through simulated environments. Twelve participants completed the second experiment; five were current smartwatch users who were recruited to investigate the effects of smartwatch experience. In this second experiment, the participants received arithmetic questions as notifications, and were required to manually select, from three choices, a correct reply to each question. Selecting a reply to an arithmetic question required deeper cognitive processing of the notification content than the earlier read aloud task, and represents the more visually and manually demanding tasks that are possible with smartwatches (e.g., responding to text messages using canned responses).

Both experiments collected self-reported data on perceived risk associated with the use of these devices in simulated driving scenarios. In addition, the second experiment collected data on the perceptions and knowledge about relevant distracted driving legislation and penalties. A study on Google Glass suggests that users of wearable technologies may be misled by some of the obvious benefits provided by the device (e.g., gaze staying in the direction of travel) and overestimate their ability to drive well while interacting with the device (He, Ellis, Choi, & Wang, 2015). With smartwatches, such overestimation may occur due to the ease of access to notifications and may lead to uncalibrated risk assessment. Previous research has found risk perception about cell phone use while driving to be correlated with actual usage (e.g., Hallett, Lambert, & Regan, 2011), but to our knowledge, no studies have examined perceived risk associated with smartwatch use.

As smartwatches have entered the consumer market only recently, guidelines and policies are not yet clear with respect to their usage while driving. For example, in Ontario, Canada, where our studies were conducted, laws forbid the use of hand-held devices while driving. Smartwatches, which are neither hand-held nor hands-free, are not currently addressed explicitly. Lack of clear legislation along with the marketed image of smartwatch use may contribute to inaccurate perception of risk associated with smartwatch use while driving, which in turn may lead to inappropriate use.

Overall, the objectives of our studies were to understand whether the unique features of a smartwatch, compared to a smartphone, mitigate or exacerbate driver distraction due to notifications, and to provide insights about drivers' perception of the risks associated with using smartwatches while driving. Both experiments were approved by the University of Toronto Research Ethics Board.

2. EXPERIMENT I - VERBAL RESPONSE TO NOTIFICATIONS

The goal of the first experiment was to explore how drivers engage with notifications on a smartwatch compared to a smartphone. Participants were asked to read aloud the notification content, and their visual-manual response was investigated. Preliminary findings from this experiment were reported in Giang, Hoekstra-Atwood, & Donmez (2014).

2.1. Participants

Six participants (3 male and 3 female) from the local community and the undergraduate and graduate population at the University of Toronto participated in this study. Participants were right handed with normal or corrected-to-normal vision and reported that they owned and regularly used a smartphone while driving. Further, all participants had a full driver's license and reported driving at least 5,000 km in the previous year. In addition to these criteria, the participants were also screened for proneness to simulator sickness.

Participant ages ranged between 22 and 31; mean (M) = 25.5 years, standard deviation (SD) = 3.33. The phone activities participants reported engaging in while driving included calling hand-held or hands-free (100% of participants), texting (67%), GPS (33%), and reading other notifications (17%). Participants had high self-reported scores for technology experience (rated out of 10, M = 9, SD = 0.63), for ability to learn how to operate new technologies (rated out of 10, M = 8.5, SD = 1.05), and for whether they thought they were a safe driver (rated out of 10, M = 7.2, SD = 0.98). Upon completion of the experiment, participants were compensated C$15.

2.2. Apparatus

The study was conducted using a NADS quarter-cab MiniSim™ driving simulator (Figure 1), which uses three 42-inch 1024x768 plasma wide-screen displays that create a 130° horizontal and 24° vertical field of view at a 48-inch viewing distance. An additional 19-inch screen integrated into the dash displays speedometer and revolution meter. The simulator uses an authentic steering wheel, column gear selector, pedals, and vehicle seat. Stereo sound of the vehicle and its surroundings is portrayed through two speakers in the front; a third speaker mounted below the driver seat simulates roadway vibrations.

The smart devices used were a Pebble Inc. Pebble smartwatch and a LG Nexus 5 Android smartphone. The Pebble is a Bluetooth connected smartwatch which receives notifications from smartphones. It uses an e-paper display with an LED backlight for use in the dark, and a vibrotactile motor which pulsates when a notification is received. The Pebble does not have a touch screen, but can be interacted with through buttons on the side of the watch. The Nexus 5 has a 4.95-inch touchscreen display.

The Pebble's default actionable notification system was used on the smartwatch but a custom messaging application was used for the Nexus 5. During the experiment, participants wore the smartwatch on their left wrist, and the smartphone was placed on a table beside the driver's seat between uses (Figure 1). The devices were removed when they were not used in a given experimental condition. Participant interactions with the devices were recorded on two cameras (720x480 at 29 fps): one provided a front view of the participant, including the participant's face, and the other provided an overhead view of the cab. During the experiment, overhead lights in the simulator room were turned off.

Figure 1. Experiment I apparatus: NADS MiniSim simulator with a Pebble smartwatch on the participant's left wrist and the Nexus 5 smartphone placed to the right of the participant

2.3. Experimental Design

The study followed a repeated measures design. The independent variable, i.e., the notification medium, had three levels: smartwatch, smartphone, and baseline with no device/notification. The presentation order of these levels was counterbalanced to control for learning effects.

2.4. Driving Scenarios

Four driving scenarios were created: three for the three experimental conditions and one for a practice drive. These scenarios made use of the same road network containing a rural environment for the first half and an urban environment for the second, and took approximately 10 minutes to complete. Lane width across the entire road network was 12 feet (3.7 m). The rural environment consisted of a two-lane highway with gravel shoulders and a posted speed limit of 50 mph (80.5 km/h). There were three straight sections separated by two curves. In the urban environment, a four-lane road, with a posted speed limit of 35 mph (56.3 km/h), connected six four-way intersections. The road was divided by a double solid line, with sidewalks as well as parked cars, buildings, and stationary pedestrians on each side. At two of the intersections, the participants were asked to make a left-hand turn. Oncoming traffic was present in both environments. Participants were instructed that the safe operation of the vehicle was their primary task.

A series of events, which were similar across different scenarios, occurred during the drive (e.g., lead vehicle braking and bicyclists entering the roadway). These events were included to increase participants' vigilance to the driving task but were not precisely timed with respect to the onset of notifications. Due to this imprecision, driving performance is not analyzed for this experiment, and notifications that occurred during and immediately after the events were also excluded from our analysis.

2.5. Notification Task

For the two device conditions, participants were asked to read aloud the notifications that were sent to the device. All notifications contained a randomly generated address consisting of three numbers and a street name (e.g., 735 Canada Drive). Participants were informed that they could defer the task until they felt it was safe to engage with the notification, but they should do so in a timely manner.

On the smartwatch, a single pulse of the watch's vibrating motor alerted the participant of an incoming notification. The address was displayed on the screen, and the backlight of the display turned on for 2 s. If there was a need, shaking the watch or pressing any of the buttons also turned the backlight on. In the smartphone condition, participants were alerted through an auditory tone. The backlight of the phone turned on with the address displayed on the preview screen and timed out after 10 s of inactivity. No such timeouts were observed as any interaction with the smartphone's interface prolonged the duration of the backlight. Participants were instructed to pick up the smartphone from the side table to read the notification and to place the phone back on the side table when they completed the verbal response to the notification.

The notifications were sent throughout the drive, with one occurring exactly every minute and additional notifications triggered by locations within the driving scenarios. The number of notifications sent varied slightly based on driving time (M = 13 per drive, SD = 1.1). Participants were told to read aloud only the latest notification if they received a new notification before they had a chance to respond to an earlier notification. As mentioned previously, some of the notifications (M = 3 per drive, SD = 1.0) occurred around driving events and were excluded from analysis.

In addition to the ones reported above, a few more notifications (M = 4 per drive, SD = 0.9) were sent to the participants, which required additional manual interactions with the device to access the address content. Participants were presented with a message to "scroll down to read the message" in the smartwatch condition, and to "press to read the message" in the smartphone condition. In the smartwatch condition, this required participants to press the side button 2-3 times to see the message on the display. In the smartphone condition, participants were required to tap the notification on the preview screen to bring up the messaging application, and then tap the notification again to bring up the message containing the address. The relevant data are not analyzed in this paper given the limited sample size combined with the imprecise timing of these notifications with respect to driving events, but a qualitative narrative of how drivers engaged with these notifications is provided. Notifications that require manual interactions were investigated in more detail in Experiment II, which also employed a more controlled driving environment.

2.6. Procedure

Participants were first given a familiarization session that introduced the Pebble smartwatch, the Nexus smartphone, and the notification tasks; they practised reading aloud the addresses until they reported being comfortable with using the devices. Participants then completed a familiarization drive in the simulator, which also served to assess simulator sickness. Afterwards, participants proceeded to the practice drive where they received notifications on both the smartwatch and the smartphone and were asked to alternate between the two to become familiar with their use while driving. The practice drive was followed by the three experimental drives. At the end of each experimental drive, participants rated the perceived riskiness of their last drive (Tsimhoni, Smith, & Green, 2003). The scale ranged between 1 and 10, with 10 representing “driving with my eyes closed; a crash is bound to occur every time I do this” and 1 representing “driving on an easy road with no traffic, pedestrians, or animals while perfectly alert”. The entire experiment took approximately 1.5 hours.

2.7. Results

The start and finish times of the verbal response to the notifications were manually recorded by the experimenter in real time. After all data were collected, glances to the devices (onsets and offsets) were manually coded from the front camera video recordings of the participants using ISO 16673: 2007 standards. Glance onsets were defined by eye movements away from the simulator screen and towards the direction of the device, and similarly offsets were defined by eye movements away from the direction of the device back towards the simulator screen. Three categories of dependent measures were evaluated: task engagement, glance characteristics, and perceived risk. The means and differences in means reported in this section are statistical model estimates.

2.7.1. Task Engagement

Time to engagement and total task engagement time per notification were used to assess task engagement. Time to engagement was defined as the duration between when the notification was sent to the device and when the participant made their first visual glance towards the device. Total task engagement time per notification was defined as the time period between this first glance and the end of the verbal response. Using R’s ‘nlme’ package (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2015), linear mixed effects models were built for these two task engagement metrics, with notification medium (smartwatch vs. smartphone) as a fixed and participant as a random factor. The unit of analysis was an individual notification. A log transformation was used for both measures to ensure that residuals were normally distributed. The means reported here are back-transformed. The smartwatch condition had a shorter time to engagement (M = 2.25 s vs. M = 3.63 s) and a longer total task engagement time (M = 4.14 s vs. M = 3.61 s) than the smartphone condition, $F(1,113) = 27.5$, $p<.0001$ and $F(1,113) = 10.0$, $p = .002$, respectively. Thus, participants were faster to engage with the smartwatch after receiving a notification, but spent more time reading aloud the notification content.

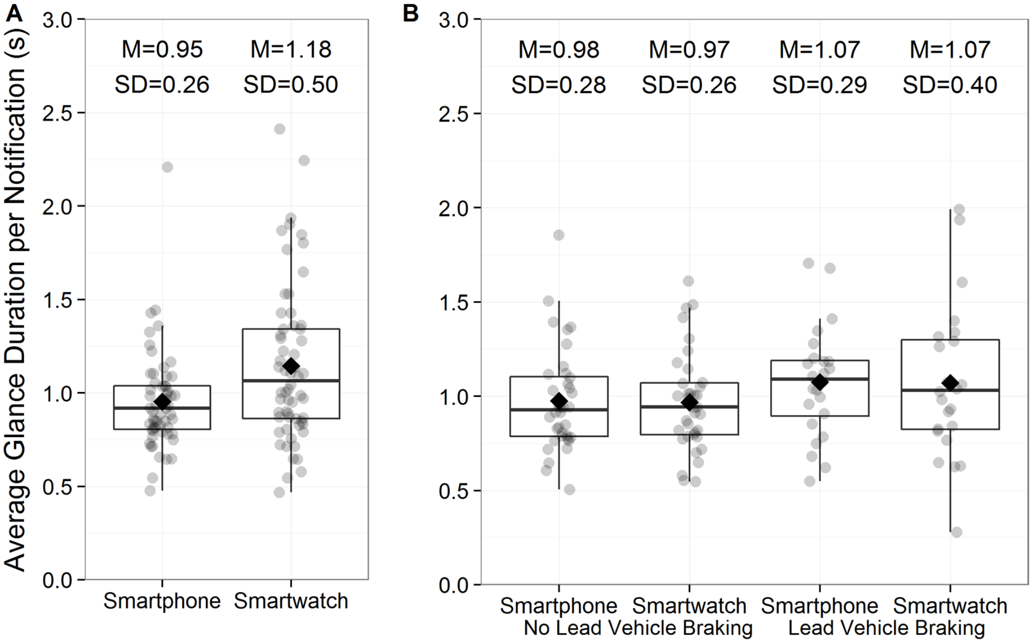
2.7.2. Glance Characteristics

Four glance metrics were analyzed: average glance duration towards the device, total number of glances and total glance duration towards the device per notification, and frequency of long glances towards the device during a drive. Using R’s ‘nlme’ package, linear mixed models were built for average glance duration, total number of glances, and total glance duration, with notification medium as a fixed and participant as a random factor. Log transformations were used to ensure normality of residuals; thus, the means reported here are back-transformed. The unit of analysis was an individual notification. The smartwatch condition had longer average glance durations (M = 1.10 s vs. M = 0.92 s, Figure 2a) and longer total glance durations per notification (M = 2.25 s vs. M = 1.80 s) than the smartphone condition, $F(1,113) = 11.9$, $p = .0008$ and $F(1,113) = 16.29$, $p<.0001$, respectively. No significant effects were observed for total number of glances ($p>.05$).

Two models were built to evaluate the total number of long device glances observed within a drive: one for glances longer than 1.6 s and one for glances longer than 2 s. Both of these glance duration thresholds have been used to study driver distraction (Reimer, Mehler, & Donmez, 2014). Off-road glances longer than 2 s elevate crash risks (Klauer et al., 2006) and 1.6 s has been reported as an upper bound for off-road glance durations observed in earlier on-road studies, with few glances going above this threshold (e.g., Sodhi, Reimer, & Llamazares, 2002). Using R's 'lme4' package (Bates, Maechler, Bolker, & Walker, 2015), a Poisson model with a log link function was built to predict the rate (per notification) of glances longer than 1.6 s with notification medium as a fixed and participant as a random factor. The number of notifications received within the drive was used as an offset variable. The smartwatch condition (median = 3 glances per drive) had 3.7 times the rate of glances longer than 1.6 s observed in the smartphone condition (median = 1 glance per drive), $\chi^2(1) = 5.44$, $p = .02$. There were very few glances longer than 2 s (smartwatch median = 0.5 glances per drive; smartphone median = 0 glances per drive), so a logistic regression model was used, but no significant differences were observed, $\chi^2(1) = 1.4$, $p = .24$.

Overall, the smartwatch induced longer average glance durations, longer total glance durations per notification, and a greater number of glances longer than 1.6 s.

Figure 2. Average glance duration per notification for Experiment I (A: across the two notification mediums) and Experiment II (B: across the two notification mediums for no lead vehicle braking and lead vehicle braking situations). For these boxplots and the following ones, the sample quartiles are presented along with the individual data points in the background and the sample means as darker black diamonds.



2.7.3. Perceived Risk

A Friedman test revealed that there was a marginally significant difference in perceived risk between the three conditions, $\chi^2(2) = 5.78$, $p = .056$. The median perceived risk value for the smartwatch, smartphone, and baseline conditions were 4.5, 5, and 3.5, respectively. Follow-up Wilcoxon signed-rank tests revealed that the baseline condition was perceived to be less risky than the smartphone condition at a statistically marginally significant level, $W = 0$, $p = .06$.

2.7.4. Notifications Requiring Manual Interaction

While not quantitatively analyzed, qualitative observations of participant interactions with the devices during notifications that required additional manual interaction with the interface revealed very different manual interaction patterns for the two devices. When participants had to scroll through the message on the smartwatch they had to use both hands: one to keep the watch in position while holding the steering wheel, and the other to manipulate the device. In contrast, participants interacted with the smartphone using only one hand, cradling the phone in their right hand while using their thumb to interact with the touch screen; their left hand could remain on the steering wheel. These differences may lead to different implications for vehicle control and driving performance. Thus, Experiment II was conducted to further investigate interactions with notifications that require a manual response and their impact on driving performance.

3. EXPERIMENT II – MANUAL RESPONSE TO NOTIFICATIONS

Experiment II focused on how drivers engage with notifications that require manual interaction with the interface (i.e., answering arithmetic questions using canned responses) during a lead vehicle following task. Visual-manual response and driving performance were investigated. Preliminary analysis from this experiment were reported in Giang, Shanti, Chen, Zhou, & Donmez (2015).

3.1. Participants

Twelve male participants between the ages of 22 and 29 (M = 23.9 years, SD = 2.27) participated in this study. Five were current smartwatch users (3 Pebble smartwatches and 2 Moto 360s), the rest were novice users with no smartwatch experience. The current users were active smartwatch users, with four out of five reporting to checking notifications daily and frequently and all five reporting to checking notifications automatically and without thinking.

The participants were recruited from the same population used in Experiment I and were screened using the same set of criteria (right handedness, normal or corrected-to-normal vision, regular smartphone usage while driving, full driver licensure, and lower likelihood of simulator sickness). Due to the difficulty in recruiting smartwatch users, the mileage criteria was relaxed in the second experiment from having driven at least 5,000 km in the previous year to having to drive at least a few days each week.

Overall, participants had high self-reported scores for technology experience (rated out of 10, M = 9.2, SD = 0.83), for ability to learn how to operate new technologies (rated out of 10, M = 9.2, SD = 1.1), and for whether they thought they were a safe driver (rated out of 10, M = 7.8, SD = 0.8). The phone activities participants reported engaging in while driving included reading text messages (83% of participants), navigation/GPS (83%), hands-free (75%) and hand-held calling (33%), sending text messages (42%), and reading other notifications (58%). The smartwatch users reported using their smartwatches while driving for reading text messages (80%), navigation (60%), and reading other notifications (60%). Participants were compensated C$25 for completion of the experiment.

3.2. Apparatus

The same apparatus was used as Experiment I (Section 2): a MiniSim driving simulator, a Pebble Inc. Pebble smartwatch, and a LG Nexus 5 Android smartphone. In addition, participants' eye movements were tracked using a head-mounted Ergoneers Dikablis Professional Eye-Tracking System with two eye cameras and a forward scene camera (Figure 3). The locations of the smartwatch and the smartphone remained the same as Experiment I. The overhead lights, however, were kept on throughout the experiment to improve the readability of the device displays and to reduce the reliance on the backlight for the Pebble smartwatch, which has an e-paper display. The custom messaging application on the smartphone was also modified to allow for the recording of data from both the smartphone and the smartwatch, including the notification content as well as the content and the timing of participant replies.

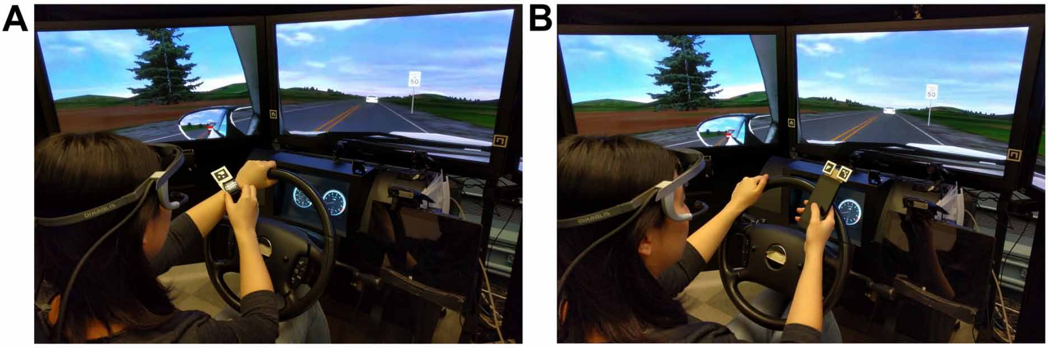
3.3. Experimental Design

The study was a mixed factorial design with notification medium (smartwatch, smartphone, and baseline) as a within-subject variable and smartwatch experience as a between-subject variable (novice vs. current smartwatch user). Experiment I employed a single 10-minute long scenario for each experimental condition. Experiment II increased the total driving time per condition to 12 minutes but broke it into three distinct 4-minute scenarios to prevent fatigue effects. The presentation order of the notification mediums was counterbalanced and the three scenarios were repeated in the same order for all participants and were blocked by notification medium (first medium: scenario 1, 2, and 3; second medium: 2, 3, and 1; third medium: 3, 1, and 2).

3.4. Driving Scenarios

The three driving scenarios were created using a driving environment similar to the rural portion of Experiment I, consisting of a straight rural two-lane highway with gravel shoulders and a posted speed limit of 50 mph (80.5 km/h). The road had a lane width of 12 feet (3.7 m) and was 2.6 miles (4.2 km). Participants followed a lead vehicle for the entire duration of the scenario. They were asked to maintain the posted speed limit and were followed by a second vehicle to encourage them to do so. The lead vehicle occasionally braked at a rate of 0.5 g at an average headway time of 2.99 s (SD = 0.96). Prior to braking, the lead vehicle smoothly adjusted its speed to achieve a headway time of 2.1 s, but this exact headway time was not always reached due to participants failing to maintain speeds close to the speed limit. Scenarios 1 and 2 contained two braking events each, and Scenario 3 had one braking event. Oncoming traffic was present in all scenarios. Participants were instructed that the safe operation of the vehicle was their primary task.

Figure 3. Experiment II apparatus: NADS MiniSim simulator with (A) Pebble smartwatch, (B) Nexus 5 smartphone

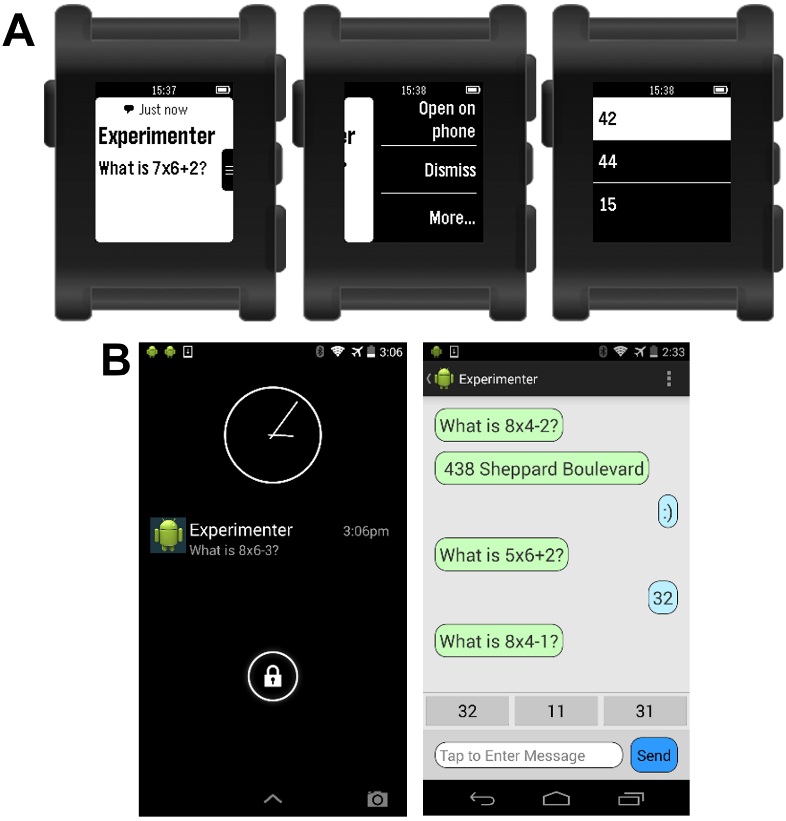


3.5. Notification Task

In each device condition, the participants received a total of five notifications that were multiple-choice arithmetic questions (3 choices). These questions involved the multiplication of two single digit numbers followed by the addition or subtraction of a single digit number (e.g., 5 x 4 – 7). Selecting a reply to an arithmetic question required deeper cognitive processing of the notification content than the earlier read aloud task, and represents the more visually and manually demanding tasks that are possible with smartwatches (e.g., responding to text messages using canned responses). Of these five notifications, two occurred immediately prior to a lead vehicle braking event (~2 s from notification alert to lead vehicle brake light onset), and three occurred during non-braking periods. As per Experiment I, participants were informed that they could defer the task until they felt it was safe to engage with the notification, but to do so in a timely manner.

The same notification alerts were used as Experiment I: a single vibrotactile pulse on the smartwatch and an auditory tone on the smartphone. The arithmetic question was displayed concurrently with the alert and the participants were required to mentally calculate and manually select an answer (Figure 4). For the smartwatch, the response included navigation though three screens via three to five side-button presses (depending on the placement of the correct answer). For the smartphone, the response included three screen taps across two screens. Participants were instructed to pick up the smartphone from the side table to read and reply to the notification and to place the phone back on the side table when they completed the notification task. Due to the limited field-of-view of the eye tracker, participants were asked to not manipulate the devices directly on the

Figure 4. Device interface and response process for Experiment II: (A) smartwatch and (B) smartphone



Experimenter
What is 7x6+2?

Experimenter
What is 8x6-3?

Experimenter
What is 8x4-2?
438 Sheppard Boulevard
:)
What is 5x6+2?
32
What is 8x4-1?

table or in their laps, as that would have placed the device outside of the view of the scene camera. Typical interactions observed are demonstrated in Figure 3.

Each device condition had two additional notifications similar to the ones used in Experiment I. These notifications contained an address consisting of three numbers and a street name (e.g., 735 Canada Drive) but did not require a read-aloud response; they were implemented to investigate notifications which require a simple visual interaction with the devices. During data collection, it appeared that while some participants may have been reading the content, some may have just been filtering these notifications by a simple confirmation that they were not an arithmetic notification requiring a reply. Due to this variability in task engagement, the data related to these notifications are not reported in this paper.

3.6. Procedure

After filling out a battery of questionnaires, the participants were familiarized with the smartwatch and the smartphone, and practiced the notification task until they were comfortable performing it using both devices. Participants then completed a familiarization drive in the simulator, which also assessed their proneness to simulator sickness. Following the familiarization drive, they completed three practice drives, one for each experimental condition (smartwatch, smartphone, and baseline). These practice drives were similar to the experimental scenarios. The participants were given the option to repeat the practice drives if they were not yet comfortable with the tasks, but no participant opted to repeat. The participants then completed the experimental scenarios, and after each experimental condition, filled out the perceived risk questionnaire used in Experiment I (described in Section 2). At the end of the experiment, participants were asked whether it is legal to use the following devices while driving in Ontario: hand-held devices, hands-free devices, and smartwatches (yes, no, or I don't know). After answering, they were informed that it was legal to use hands-free but not hand-held devices, and that the legislation did not explicitly classify smartwatches as hands-free or hand-held. The participants were then asked whether legislation about hand-held devices should also apply to smartwatches (yes, no, or undecided), and if so, how strict the penalties should be (less than/equal to/more than the current penalties for hand-held devices, or undecided). The entire experiment took approximately 2 hours.

3.7. Results

After data collection, glances to the devices were coded following ISO 16673: 2007 standards in the D-Lab software that comes with the eye-tracker used in the study. Similar to Experiment I, dependent measures for task engagement, glance characteristics, and perceived risk were analyzed. In addition, driving performance and responses to legislation-related questions were also evaluated. Where relevant, notifications that occurred immediately prior to lead vehicle braking events were separately analyzed. The means and differences in means reported in this section are statistical model estimates.

3.7.1. Task Engagement

The eye-tracking system used in Experiment II was not synchronized with the simulator and thus for this experiment, it was not possible to capture the task engagement metrics used in Experiment I, which relied on glances to the device. Instead, total task duration per notification was used, which was calculated as the duration from when the notification was sent to the device until the participant selected a reply. Two linear mixed models (lead vehicle braking present vs. not) were built with notification medium (smartwatch vs. smartphone), smartwatch experience (novice vs. current user), and their interaction as fixed factors and participant as a random factor. Of the 120 total notifications, only five had incorrect answers. These cases, as well as one additional case with an exceptionally long total task duration (25 s) were removed from the analysis of this dependent variable. The unit of analysis was an individual notification.

For lead vehicle braking cases, both the interaction effect, $F(1,30) = 3.7$, $p = .06$, and the main effect of smartwatch experience, $F(1,10) = 4.4$, $p = .06$, were marginally significant. Pairwise comparisons computed using R's 'multcomp' package (Hothorn, Bretz, Westfall, & Heiberger, 2008) showed that the total task duration with the smartphone was shorter for novice smartwatch users (M = 9.3 s) than for current users (M = 13.3 s), $Z = -2.1$, $p = .04$; a similar effect was not observed for the smartwatch condition (novice user M = 10.9 s; current user M = 11.6 s). For no lead vehicle braking cases, the interaction effect was significant, $F(1,56) = 4.3$, $p = .04$. Novice smartwatch users had task durations that were longer on the smartwatch (M = 9.6 s) than on the smartphone (M = 8.3 s), $Z = -2.3$, $p = .02$; a similar effect was not observed for current smartwatch users (smartwatch condition M = 9.2 s; smartphone condition M = 9.7 s).

Taken together, these results suggest that total task duration was largely similar between novice and current smartwatch users, and between smartwatch and smartphone conditions. However, the novice users tended to complete the task faster on the smartphone than they did on the smartwatch during non-braking periods. Further, when there was a lead vehicle braking event, the novice smartwatch users were faster than the current users in completing the task on the smartphone.

3.7.2. Glance Characteristics

The same four glance metrics used in Experiment I were used in Experiment II as well: average glance duration towards the device, total number of glances and total glance duration towards the device per notification, and frequency of long glances towards the device during a drive. Linear mixed models were built for the first three variables, with notification medium (smartwatch vs. smartphone), smartwatch experience (novice vs. current users), and their interaction as fixed factors and participant as a random factor. The unit of analysis was an individual notification. Interactions were not significant and were removed from the final models.

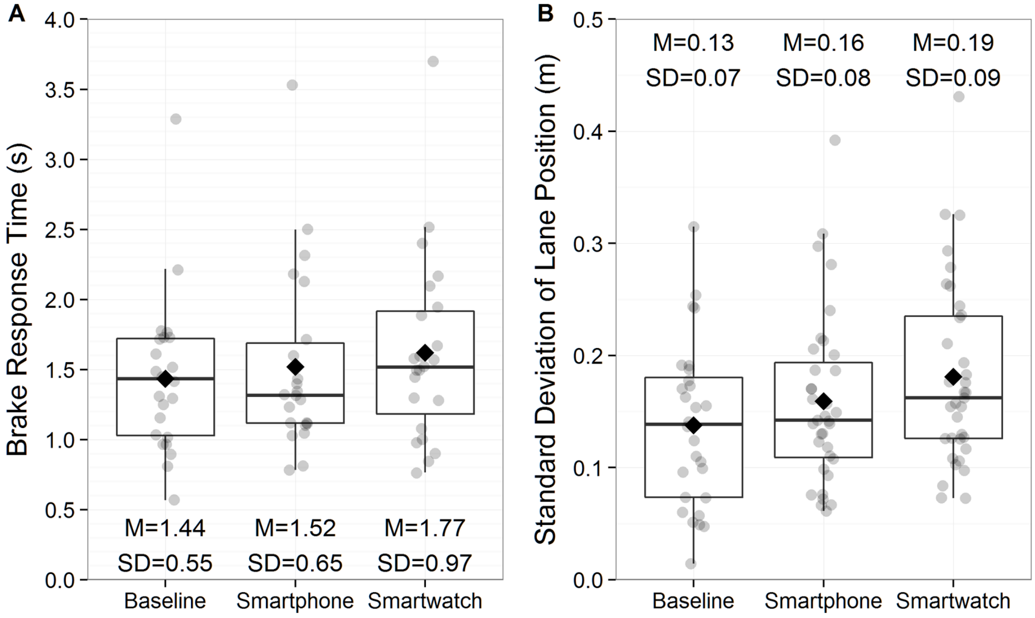
The smartwatch condition produced more glances per notification (braking event M = 4.9; no braking event M = 5.2) than the smartphone condition (braking event M = 3.8; no braking event M = 4.2), both when there was a lead vehicle braking event, $F(1,35) = 8.35$, $p = .007$, and when there was none, $F(1,59) = 9.18$, $p = .004$. The smartwatch condition also produced longer total glance durations per notification (braking event M = 5.03 s; no braking event M = 4.95 s) than the smartphone condition (braking event M = 3.97 s; no braking event M = 3.98 s), both when there was a lead vehicle braking event, $F(1,35) = 6.45$, $p = .02$, and when there was none, $F(1,59) = 9.49$, $p = .003$. There were no significant differences for average glance duration ($p>.05$, Figure 2b). Thus, while responding to notifications, participants tended to have more glances towards the smartwatch than the smartphone, and this difference led to longer total time spent looking at the smartwatch. Poisson and logistic regression models were built for glances longer than 1.6 s and for glances longer than 2 s, respectively. However, no significant effects were observed ($p>.05$).

3.7.3. Driving Performance

Driving performance (Figure 5) was assessed by brake reaction time (BRT) in response to lead vehicle braking and by standard deviation of lane position (SDLP) during non-braking periods.

BRT was defined as the time from lead vehicle brake light onset to the participant applying force on the brake pedal. As mentioned earlier, for the device conditions, there were two braking events that happened immediately following a notification alert. The corresponding braking events (i.e., events that happened in the same road segment) from the baseline condition were used for comparison. A linear mixed model was built, with notification medium (smartwatch, smartphone, baseline), smartwatch experience (novice vs. current users), and their interaction as fixed factors and participant as a random factor. The unit of analysis was an individual braking event. To control for the variability in headway time and participant speed at the time of lead vehicle brake light onset, these two variables were used

Figure 5. (A) Brake response time and (B) Standard deviation of lane position across the three notification mediums



as covariates in the model. Two cases with extremely short BRTs (<250 ms) were excluded. A log transformation was used to ensure normality of the residuals. The interaction term and the main effect of experience were not significant and were removed from the final model. Notification medium was found to be significant, $F(2,54) = 3.45$, $p = .04$. Pairwise comparisons showed that BRTs in the smartwatch condition were 1.23 times the length of those in the baseline condition, $Z = 2.54$, $p = .01$, and 1.17 times the length of those found in the smartphone condition, $Z = 1.86$, $p = .06$; the latter finding was marginally significant. No difference was found between the smartphone and the baseline conditions, $Z = 0.59$, $p = .55$. Thus, receiving a notification on the smartwatch appeared to delay BRTs compared to receiving one on the smartphone as well as to baseline driving.

SDLP was calculated based on the lateral position relative to the lane center. For device conditions, SDLP was calculated over the period from when the notification was sent until 5 s after the participant made a reply on the device; the additional 5 s after the reply were included to capture potential vehicle control corrections made post-interaction. Corresponding road sections from the baseline condition were used as a comparison. As mentioned previously, notifications that occurred prior to lead vehicle braking were excluded from SDLP analysis. A linear mixed model was built, with notification medium (smartwatch, smartphone, baseline), smartwatch experience (novice vs. current users), and their interaction as fixed factors and participant as a random factor. The unit of analysis was an individual notification. A log transformation was used to ensure the normality of residuals. The interaction term and the main effect of experience were not significant and were removed from the final model. Notification medium was found to be significant, $F(2, 94) = 10.5$, $p<.0001$. Pairwise comparisons showed that SDLP in the smartwatch, $Z = 4.49$, $p<.0001$, and the smartphone conditions, $Z = 3.07$, $p = .002$, were 1.56 and 1.35 times those found in similar driving segments in the baseline condition. No difference was found between the smartwatch and the smartphone conditions, $Z = 1.42$, $p = .16$. Overall, participants appeared to have poorer lateral vehicle control while interacting with the two devices than while not.

3.7.4. Perceived Risk and Distracted Driving Legislation Knowledge

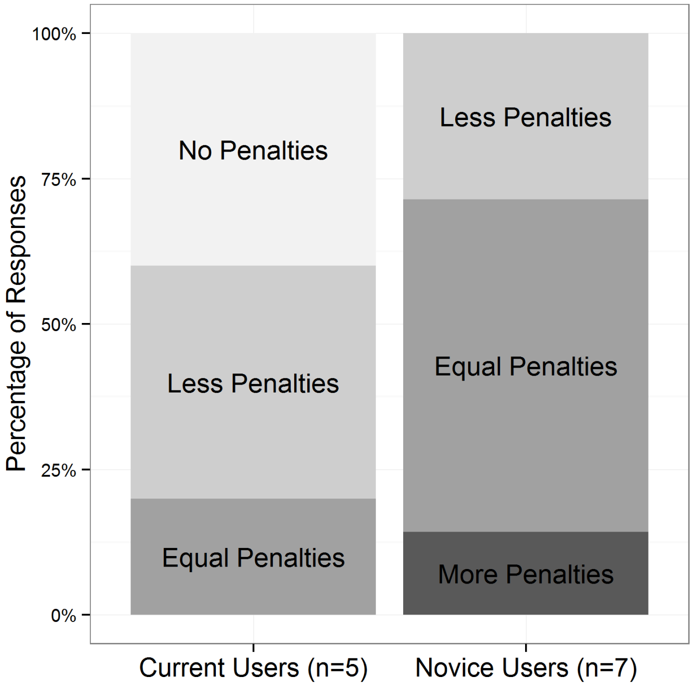
A Friedman test revealed a significant effect of notification medium on perceived risk, $\chi^2(2) = 16.2$, $p<.001$ (smartwatch median = 4.5; smartphone median = 4; baseline median = 1). Follow-up Wilcoxon signed-rank tests showed significant differences in risk perception between smartwatch and baseline, $W = 0$, $p = .002$, and between smartphone and baseline conditions, $W = 0$, $p = .004$, with no significant difference between smartwatch and smartphone conditions, $W = 16$, $p = .47$.

Eleven out of the 12 participants (92%) knew that it was illegal to operate a hand-held device, and all 12 participants knew that it was legal to operate hands-free devices while driving. However, almost all participants (11), in particular all current smartwatch users, were unaware whether it was legal to use a smartwatch while driving. When asked about future penalties on smartwatch usage while driving, 11 said that smartwatches should get penalties equal to or less than smartphones. Current smartwatch users appeared to be more lenient with their penalty judgements than novice smartwatch users (Figure 6).

4. DISCUSSION

The two studies described in this paper examined how drivers interact with smartwatch and smartphone notifications, and how such interactions impact driving performance. The first study looked at a notification task that focused on visual information retrieval, mimicking a usage scenario such as reading a text message or the next step in turn-by-turn navigation. A read aloud response was employed to ensure that information retrieval was completed. The second study examined a usage scenario where the information retrieval was followed by a manual response, simulating real-world tasks such as responding to incoming text messages with a canned response or making a selection

Figure 6. Ratings for possible penalties for smartwatch use while driving in comparison to penalties for smartphone use



based on suggestions provided by a mobile application. This latter task had stronger cognitive and manual components compared to the one used in the first experiment.

The first experiment provided insights into how quickly drivers visually engage with notifications; this measure could not be assessed in Experiment II due to apparatus setup limitations. The results showed that participants engaged faster with the smartwatch than the smartphone that was at a larger visual eccentricity from the forward road scene. Further, the smartwatch was in a more accessible location (i.e., on the wrist vs. on the table beside the driver's seat) and required less effort to access the notification content (e.g., rotating the wrist so that display is visible vs. turning the head to see the phone or reaching to pick it up to see the display). The smartwatch's vibrotactile alert might have also provided a more salient cue than the smartphone's auditory alert. Vibrotactile alerts have been found to be useful for conveying and manipulating sense of urgency (e.g., Baldwin et al., 2012; Baldwin & Lewis, 2014) and may have a stronger effect interrupting an on-going task than auditory or visual signals (Lu, Wickens, Sarter, & Sebok, 2011; Sarter, 2006). For example, tactile warnings have been found to generate faster reaction times than auditory warnings for preventing rear-end collisions in driving simulator studies (e.g., Mohebbi, Gray, & Tan, 2009; Scott & Gray, 2008). While both device accessibility and notification alert saliency might have contributed to faster engagement times, differentiating the effects of these potential factors remains a direction for future work.

Once visually engaged with the device, the total time participants spent looking towards the device during the notification period was longer with the smartwatch compared to the smartphone, a finding that was also observed in Experiment II. In the first experiment, this difference was primarily the consequence of longer average glance durations towards the device, which may be attributed to participants having difficulty reading the smartwatch's smaller display. In the second experiment, average glance durations were not different between the two devices, but there were more glances towards the smartwatch in a given notification period leading to the longer total glance durations observed. The presence of overhead lighting in the second experiment and the absence of it in the first one might have led to the difference between the two. In lower levels of ambient light, the Pebble users have to rely on its backlight, which has a 2 s timeout period. Potential timeouts coupled with the potentially degraded legibility with the backlight may have led to the longer glances observed in Experiment I. Further research is needed to investigate the use of these devices during day vs. nighttime driving conditions.

Another potential explanation for this difference in average glance duration findings between the two experiments is the differences in number of screens used in each device condition. Both devices used one screen for the read-aloud task of Experiment I. For Experiment II, the smartwatch used three screens, while the smartphone used two screens and presented both the question and the choices on a single screen (Figure 4). The task being split into multiple screens in the smartwatch condition may have provided natural breakpoints to enable the drivers to glance back at the road, mitigating the longer average glance durations observed with the smartwatch in Experiment I (as evidenced by a lack of difference between the two device conditions in Experiment II and by the larger number of glances per notification observed in Experiment II but not in Experiment I). These kinds of secondary task boundaries have been shown to influence driver glance and task switching behavior in other simulator studies (Lee, Gibson, & Lee, 2015).

Despite the fact that average glance durations were not different between the two devices in Experiment II, as discussed above, the frequency of glances were larger for smartwatches leading to a longer total time spent looking at the smartwatch for a given notification. Further, the smartwatch task required more manual interactions with the interface. These differences in visual/manual demand may have generated the slower brake response times observed in the smartwatch condition compared to the smartphone condition. No differences in brake response time were observed for either device in comparison to baseline driving. However, both devices degraded lateral vehicle control. Other interaction methods, such as voice commands, may mitigate the visual/manual demands imposed by both devices (e.g., Barón & Green, 2006; Graham & Carter, 2000; Ranney, Harbluk, & Noy, 2005;

Samost et al., 2015). Further, since smartwatches are designed to help with filtering information, they may be restricted to providing simple non-interactive information, requiring users to check their smartphones (hopefully during a safe time) when more action is required. Future research should investigate the effects of such visual-only interactions on driving performance.

In either experiment, there was no evidence that perceived risk was different between smartphone and smartwatch use while driving. Experiment II, which had a much simpler driving scenario, found the baseline condition to be perceived as less risky than the device conditions. The questionnaire results from Experiment II also showed that all participants but one was unaware of current legislation pertaining to the use of smartwatches while driving. Furthermore, all but one stated that smartwatches should have penalties less than or equal to smartphones. In particular, current smartwatch users appeared to be more lenient with their penalty judgements, even though no major differences in glance behaviour and driving performance were found between them and novice users. Along with the driving-related results, these questionnaire responses suggest that participants may have underestimated the risk associated with smartwatches. Further investigation of risk perception and knowledge of legislation around smartwatches in particular, and wearables in general, using a large-scale survey would add to a better understanding of the social and psychological aspects of wearable-device use while driving.

The small sample size in both studies is a significant limitation of this research. For example, with only five current smartwatch users in the second study, it was difficult to investigate smartwatch experience. Additionally, given that smartwatches are still relatively new, even the current users recruited for this study may not have demonstrated expert behaviors. As smartwatches are becoming increasingly popular, future studies can recruit a larger number of experienced smartwatch users, and investigate how experience and habitual use may lead to different use patterns while driving. Further, a controlled interaction paradigm was employed in this experiment. That is, the participants were asked to respond to the notifications as quickly as possible when deemed safe to do so. Future research should explore more natural interactions where the participants can pace their interaction as they would normally do in the real world. Using this latter interaction paradigm would enable researchers to assess the role of voluntary vs. involuntary automatic responses in notification engagement in smartwatch use.

Finally, the results we observed using the Pebble smartwatch and the Nexus 5 smartphone may not necessarily generalize to other smartwatches and smartphones. How users interact with smart devices can depend on how their interfaces are designed and the functionalities they provide. Digital and physical characteristics of other smartwatches and smartphones may lead to different use patterns and glance behaviours that should be investigated in the future.

ACKNOWLEDGMENT

We acknowledge the support of Liberty Hoekstra-Atwood, Ines Shanti, Alex Zhou, Sabrina Ralph, Carl Lam, Matt Agostini, and Kenny Kim from the University of Toronto, for their help in collecting data and programming the experimental apparatus. We also gratefully acknowledge the support of the Natural Sciences Research Council of Canada (Discovery Grant, #402275-2011).

REFERENCES

Baldwin, C. L., Eisert, J. L., Garcia, A., Lewis, B., Pratt, S. M., & Gonzalez, C. (2012). Multimodal urgency coding: Auditory, visual, and tactile parameters and their impact on perceived urgency. *Work (Reading, Mass.)*, *41*, 3586–3591. PMID:22317267

Baldwin, C. L., & Lewis, B. A. (2014). Perceived urgency mapping across modalities within a driving context. *Applied Ergonomics*, *45*(5), 1270–1277. doi:10.1016/j.apergo.2013.05.002 PMID:23910716

Barón, A., & Green, P. (2006). *Safety and Usability of Speech Interfaces for In-vehicle Tasks While Driving: A Brief Literature Review (No. UMTRI-2006-5)*. Ann Arbor, MI: University of Michigan, Transportation Research Institute.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.17. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Bayer, J. B., & Campbell, S. W. (2012). Texting while driving on automatic: Considering the frequency-independent side of habit. *Computers in Human Behavior*, *28*(6), 2083–2090. doi:10.1016/j.chb.2012.06.012

Danova, T. (2015). The Wearables Report: Growth Trends, Consumer Attitudes, and Why Smartwatches Will Dominate. *Business Insider*. Retrieved from http://www.businessinsider.com/

Fitch, G. M., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., & Olson, R. L. … Dingus, T. A. (2013). The Impact of Hand-Held and Hands-Free Cell Phone Use on Driving Performance and Safety-Critical Event Risk (No. DOT HS 811 757). Washington, DC: National Highway Traffic Safety Administration.

Giang, W. C., Shanti, I., Chen, H.-Y. W., Zhou, A., & Donmez, B. (2015). Smartwatches vs. smartphones: A preliminary report of driver behavior and perceived risk while responding to notifications. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,*Nottingham, UK (pp. 154–161). doi:10.1145/2799250.2799282

Giang, W. C. W., Hoekstra-Atwood, L., & Donmez, B. (2014). Driver engagement in notifications: A comparison of visual-manual interaction between smartwatches and smartphones. *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting,*Chicago, IL (pp. 2161–2165). doi:10.1177/1541931214581454

Graham, R., & Carter, C. (2000). Comparison of speech input and manual control of in-car devices while on the move. *Personal Technologies*, *4*(2-3), 155–164. doi:10.1007/BF01324122

Hallett, C., Lambert, A., & Regan, M. A. (2011). Cell phone conversing while driving in New Zealand: Prevalence, risk perception and legislation. *Accident Analysis and Prevention*, *43*(3), 862–869. doi:10.1016/j.aap.2010.11.006 PMID:21376877

He, J., Ellis, J., Choi, W., & Wang, P. (2015). Driving while reading using Google glass versus using a smart phone: Which is more distracting to driving performance?*Proceedings of the 8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design,*Salt Lake City, Utah (pp. 281–287).

Hoekstra-Atwood, L. (2015). *Driving Under Involuntary and Voluntary Distraction: Individual Differences and Effects on Driving Performance* [Unpublished Master's Thesis]. University of Toronto, Toronto, Canada.

Hoekstra-Atwood, L., Chen, H.-Y. W., Giang, W. C. W., & Donmez, B. (2014). Measuring inhibitory control in driver distraction. *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,*Seattle, WA (pp. 1–4).

Hothorn, T., Bretz, F., Westfall, P., & Heiberger, R. (2008). *Multcomp: Simultaneous Inference for General Linear Hypotheses. R Package Version 1.0-3*.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The Impact of Driver Inattention on Near-crash/crash Risk: An Analysis Using the 100-car Naturalistic Driving Study Data (No. HS-810 594)*. Washington, DC: National Highway Traffic Safety Administration.

Lee, J. Y., Gibson, M., & Lee, J. D. (2015). Secondary task boundaries influence drivers' glance durations. *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,*Nottingham, UK (pp. 273–280). doi:10.1145/2799250.2799269

Lu, S. A., Wickens, C. D., Sarter, N. B., & Sebok, A. (2011). Informing the design of multimodal displays: A meta-analysis of empirical studies comparing auditory and tactile interruptions. *Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting,*Las Vegas, NV (pp. 1170–1174). doi:10.1177/1071181311551244

Marulanda, S., Chen, H.-Y. W., & Donmez, B. (2015). Capturing voluntary, involuntary, and habitual components of driver distraction in a self-reported questionnaire. *Proceedings of the 8th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design,*Salt Lake City, Utah (pp. 358–364).

Mohebbi, R., Gray, R., & Tan, H. Z. (2009). Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *51*(1), 102–110. doi:10.1177/0018720809333517 PMID:19634313

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, *16*(1), 105–114. doi:10.1007/s00779-011-0412-2

Pickrell, T. M. (2015). *Driver Electronic Device Use in 2013 (No. DOT HS 812 114)*. Washington, DC: National Highway Traffic Safety Administration.

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: The Journal of the Human Factors and Ergonomics Society*, *47*(2), 439–454. doi:10.1518/0018720054679515 PMID:16170949

Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis and Prevention*, *43*(5), 1771–1781. doi:10.1016/j.aap.2011.04.008 PMID:21658505

Reimer, B., Mehler, B., & Donmez, B. (2014). A study of young adults examining phone dialing while driving using a touchscreen vs. a button style flip-phone. *Transportation Research Part F: Traffic Psychology and Behaviour*, *23*, 57–68. doi:10.1016/j.trf.2013.12.017

Samost, A., Perlman, D., Domel, A. G., Reimer, B., Mehler, B., & Mehler, A., … McWilliams, T. (2015). Comparing the relative impact of smartwatch and smartphone use while driving on workload, attention, and driving performance. *Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting,* Los Angelas, CA (pp. 1602–1606). doi:10.1177/1541931215591347

Sarter, N. B. (2006). Multimodal information presentation: Design guidance and research challenges. *International Journal of Industrial Ergonomics*, *36*(5), 439–445. doi:10.1016/j.ergon.2006.01.007

Scott, J. J., & Gray, R. (2008). A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(2), 264–275. doi:10.1518/001872008X250674 PMID:18516837

Sodhi, M., Reimer, B., & Llamazares, I. (2002). Glance analysis of driver eye movements to evaluate distraction. *Behavior Research Methods*, *34*(4), 529–538. doi:10.3758/BF03195482 PMID:12564557

Tsimhoni, O., Smith, D., & Green, P. (2003). *On-the-road Assessment of Driving Workload and Risk to Support the Development of an Information Manager (No. UMTRI-2003-08)*. Ann Arbor, MI: University of Michigan, Transportation Research Institute.

Walsh, S. P., White, K. M., & Young, R. (2010). Needing to connect: The effect of self and others on young peoples involvement with their mobile phones. *Australian Journal of Psychology*, *62*(4), 194–203. doi:10.1080/00049530903567229

Wayne C. W. Giang is a PhD candidate in the Human Factors & Applied Statistics Laboratory, at the University of Toronto, Department of Mechanical & Industrial Engineering. He received his BASc (2009) and MASc (2012) in systems design engineering from the University of Waterloo. His current dissertation research examines how individuals use uncertainty visualizations for supporting dynamic decision making in medical dispatch. His research interests also include wearable devices, shared digital displays, and multimodal perception and interface design.

Huei-Yen Winnie Chen is a postdoctoral fellow at the Department of Mechanical & Industrial Engineering, University of Toronto, as part of the Human Factors & Applied Statistics Laboratory. She received her PhD in human factors engineering from the University of Toronto in 2013, her MASc (2007) and BASc (2005) in systems design engineering from the University of Waterloo. Her dissertation examined visual information processing using the visual occlusion technique, modeling how expert operators' visual sampling behaviour may be guided by their own predictions and past information. She has worked on projects in multiple domains, including transportation, healthcare, defence, and aerospace. Her current research examines driver distraction, driver state monitoring and feedback strategies, wearables and driving, and social psychological facilitators of risky behaviours.

Birsen Donmez is an Associate Professor at the University of Toronto, Department of Mechanical & Industrial Engineering and is the Canada Research Chair in Human Factors and Transportation. She received her BS in Mechanical Engineering from Bogazici University in 2001, her MS (2004) and PhD (2007) in industrial engineering, and her MS in statistics (2007) from the University of Iowa. Before joining the University of Toronto, she spent two years as a postdoctoral associate at the Massachusetts Institute of Technology. Her research focuses on operator attention in multitask activities, decision support under uncertainty, and human automation interaction, with applications in surface transportation, healthcare, and unmanned vehicle operations. Her selected honors include the inaugural Stephanie Binder Young Professional Award from the HFES Surface Transportation Technical Group and an Ontario Early Researcher Award. She serves on multiple Transportation Research Board committees and as an associate editor for IEEE Transactions on Human-Machine Systems.