Effects of Age and Smartphone Experience on Driver Behavior during Address Entry: A Comparison between a Samsung Galaxy and Apple iPhone

Thomas McWilliams, Bryan Reimer*, Bruce Mehler, Jonathan Dobres & Joseph F. Coughlin
MIT AgeLab & New England University Transportation Center
77 Massachusetts Avenue, E40-29 Cambridge, MA 02139
*Corresponding author: reimer@mit.edu - (617) 452-2177

ABSTRACT
A Samsung Galaxy S4 and Apple iPhone 5s were compared in a driving simulator where participants performed visual-manual and auditory-vocal address entry tasks. Auditory-vocal tasks were associated with shorter task times, fewer off-road glances, lower workload ratings, and reduced impact on vehicle performance. Primarily nominal differences were found between devices. Older participants had more difficulty performing tasks across both modalities, and difficulties were amplified for visual-manual tasks. A pattern of performance advantages were found when participants used the same operating system as the one they personally owned, including faster task completion, fewer off-road glances, and lower reported workload coupled with lower impact on lane variability. One's familiarity with a smartphone may, to some extent, impact the apparent level of demand associated with its use while driving. These findings highlight a need for a more comprehensive understanding of how technology training can aid in minimizing interface demands.

Author Keywords
Automotive human machine interface; Smartphone use while driving; Voice interface; Speech interface; Distraction; Driver safety; Workload.

ACM Classification Keywords
H.1.2 [User/Machine Systems]: Human Factors; Human Information Processing; H.5.2 [User Interfaces]: Ergonomics; Interaction styles; Voice I/O

INTRODUCTION
Driver interactions with smartphones are generally complex, placing considerable attention on the safety concerns of use while driving. Research has shown that different types of devices have differential draws on driver attention [7, 2]. Other work investigating the modality of interactions suggests that voice interfaces may be less demanding than visual-manual equivalents [3, 8].

Younger drivers are more likely than older drivers to report using phones while driving [11]. Therefore, younger drivers are also likely have a greater degree of experience using smartphones while driving. Older drivers, despite their generally greater driving experience, have difficulty controlling the vehicle while using smartphones for tasks such as texting [9]. It is unclear to what degree these differences arise from drivers' specific experience with these technologies, or general age-related declines in physical and cognitive function.

No published data could be identified on the relative differences in demands that today's top selling phones place on drivers. Given that phone use of any kind while driving can disrupt attentional focus towards the road, it is important to be mindful of the relative costs and benefits of different devices operating across visual-manual and auditory-vocal interface modalities.

This study examines drivers' use of a Samsung Galaxy and Apple iPhone for address entry. The two smartphones have similar ways of presenting navigational functions, but rely on different hardware and software implementations. The purpose of this study is three-fold. First, we aim to assess the degree to which there are substantive differences in demand observed between the devices across both the visual-manual and auditory-vocal input modalities. Second, we aim to assess if previous observations of reduced demand associated with auditory-vocal interaction carry forward to both smartphone implementations. We subsequently assess if the effects of age on performance carry over as well. Finally, we aim to explore how the operating system (OS) for a participant's primary personal phone (i.e. a proxy for experience with either operating system) impacts observed demand with each smartphone.

METHODS

Participants
Sixty-three participants between the ages of 20-29 or 55-69 were recruited. All reported having a valid driver's licenses for 3+ years, driving on average one or more days a week, and being in reasonably good health for their age. Twenty-

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
Copyright is held by the owner/author(s)..
AutomotiveUI '15, September 01-03, 2015, Nottingham, United Kingdom
© 2015 ACM. ISBN 978-1-4503-3736-6/15/09…$15.00
DOI: http://dx.doi.org/10.1145/2799250.2799275

three participants were excluded from the analysis (5 pilot testing, 6 technical issues, 1 drowsiness, and 11 [all from the 55-69 age group] task performance difficulty). Participants excluded due to performance difficulty were unable to complete many or all tasks without heavy guidance, or voluntarily withdrew due to discomfort performing the tasks while driving. The final sample consisted of 40 participants balanced equally across age group and gender: 20-29 (M=24.6, SD=2.8), 55-69 (M=61.6, SD=3.4). Age did not differ significantly by gender ($F(1,38)=.033$, $p=.86$).

Apparatus

This study was conducted in a fixed base driving simulator consisting of a full cab Volkswagen New Beetle with a front projection screen providing an approximate 40 degree view of a virtual two lane rural road. The posted speed limit was 50 mph. Graphical updates were generated at a minimum frame rate of 20 Hz using STISIM Drive version 2.08.02 (Systems Technology, Inc., Hawthorne, CA) in response to driver's interactions with the controls. Correspondence between behaviors observed in the simulator and actual driving scenario has been established through previous research (see [5]). Two cameras were mounted in the vehicle; one on the dashboard recording the driver's face (observing gaze) and the other over the right shoulder (observing interactions with the smartphones).

Devices

Participants performed visual-manual and auditory-vocal address entry tasks on two different smartphones: a Samsung Galaxy S4 running Android 4.4 with Google Maps version 8.1.1, and an Apple iPhone 5s running iOS 8.1.2 with Apple Maps version 2.0. For vocal input participants used Samsung's S-Voice and Apple's Siri interfaces. Manual and vocal task workflows for both devices are shown in Table 1. During vocal interactions, participants were instructed to speak each digit of the address individually; for example "one-two-three" instead of "one-hundred-twenty-three". Both devices were "free-floating" in the cab and could be held in either hand.

The iPhone and Galaxy measured 4.87x2.31 in and 5.38x2.75 in respectively. The application icons appeared as similar sizes and were located on the middle of the right side of the screen. Both used sans-serif typefaces with similar design characteristics; Helvetica Neue for iOS and Roboto for Android. The Galaxy keyboard contained numbers on the top row, while the iPhone required users to access numbers by bringing up a separate keyboard.

Procedure

After obtaining informed consent and verifying eligibility, a questionnaire related to cell phone ownership and experience was completed. Participants were then instructed to enter the simulator and adjust the seat and steering wheel to their comfort. Prerecorded audio instructions were played introducing the participant to the simulator. Participants were first instructed to drive the simulator for two miles for familiarization.

The experiment consisted of four experimental blocks (2 - smartphone type x 2 - input method). Phone order was counterbalanced between subjects, and input method was randomized for each phone. Participants performed both input method blocks consecutively with each smartphone. Participants were instructed on the operation of the two input methods prior to the first task. Once participants were capable of performing tasks without assistance from the researcher, they were instructed to begin driving. Each block began with 180s of single-task driving and two practice address entry tasks. The evaluation portion of each block involved three address entry tasks (e.g. navigate to 177 Massachusetts Avenue, Cambridge MA). Ninety seconds of single-task driving separated each task. At the end of each block, participants were asked to give a numeric global workload rating for the tasks they had just performed (0 low; 10 high) see [8].

	Galaxy		**iPhone**	
	Manual	Vocal	Manual	Vocal
1	Tap Home button to wake up phone		Tap home button then slide screen to the right to unlock	
2	Open Google Maps	Double Tap Home button	Open Apple Maps	Hold down Home button
3	Tap search bar	Speak "Navigate to…"	Tap search bar	Speak "Navigate to…"
4	Type address, select when suggested		Type address, select when	
5	Select car icon (bottom of screen)		Select car icon (middle of screen)	
6	Select 'Start Navigation" (middle of screen)		Select "Start" (bottom of screen)	

Table 1. Steps for visual-manual and auditory-vocal tasks.

Data Reduction and Analysis

Participant glance data were manually coded by two independent coders, and mediated by a third [7]. Measures recorded during the three repetitions of each phone/input condition were averaged for statistical analysis, which employed a 2x2 repeated-measure ANOVA performed in R.

RESULTS

Of the 40 participants, 20 reported currently owning a smartphone with Android OS (11 in the younger group) and 17 owning an iPhone (9 in the younger group). Three older participants reported not owning a smartphone. Data from these three participants were excluded from analyses involving personal phone ownership. All but five participants reported primarily using manual input when entering an address into a navigation application. Two of these participants (1 younger) primarily used iOS; the remaining three were older.

Older participants had significantly longer task completion times (F(1,38)=60.20, p< .001). As illustrated in Figure 1, this effect was largely driven by the manual input modality, where the older group exhibited significantly longer task completion times for manual input than vocal (F (1,19) = 33.63, p < .001). No significant differences were observed between the two input modalities among the younger group (F(1,19)=2.0, p=.17). As detailed in Table 2, participants exhibited nominally shorter input times while using the smartphone with the same OS as their primary phone.

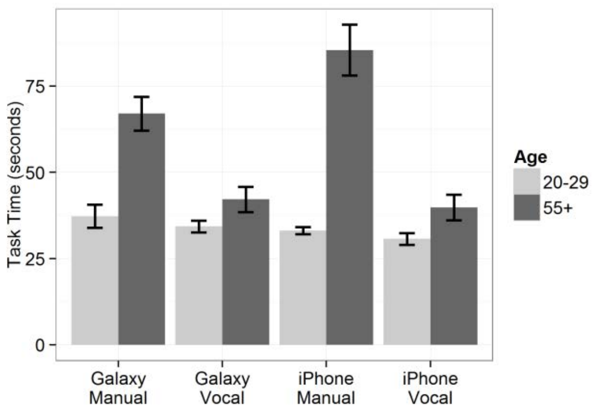

Figure 1. Task completion time. Black bars represent SEM.

	Type of OS on primary phone			
	Android		iOS\iPhone	
Phone Used	Galaxy	iPhone	Galaxy	iPhone
Task Time (s)	42.29	50.18	48.10	43.34
Workload Rating*	5.15	6.48	5.66	5.60
Off Road Glance Time (s)*	17.48	24.37	23.42	19.41
SDLP (m)*	.84	.99	.84	.77
% change in velocity (m/s)	-1.77	-0.40	-1.54	-1.66

Table 2. Comparison of Android and iPhone users using both phone types. * denotes significant effect of using personal OS.

Figure 2 illustrates workload ratings across age group, phone, and modality. Ratings were significantly higher for manual input than vocal (F(1,39)=76.22, p <.001) with older participants reporting significantly higher workload ratings for all tasks (F(1,38)=14.18, p<.001). There was a significant difference in workload ratings between the two devices (F(1,39)=11.67, P=.001). This effect is likely due to Android users reporting higher workload ratings for the iPhone compared to the Galaxy, while there was only a modest difference in rating between devices for iOS users (see Table 2). Participants provided lower workload ratings for their personal device types (F(1,36)=4.3, p=.04).

Eye glance data (Figure 3) indicates that younger participants spent less time glancing off-road (F(1,38)=40.52, p<.001). The effect was driven by longer total off-road glance times for manual input tasks (F(1,39)=59.27, p<.001). Participants spent a smaller amount of time glancing off road when using their own phone type (F(1,36)=6.68, p=.013) (Table 2).

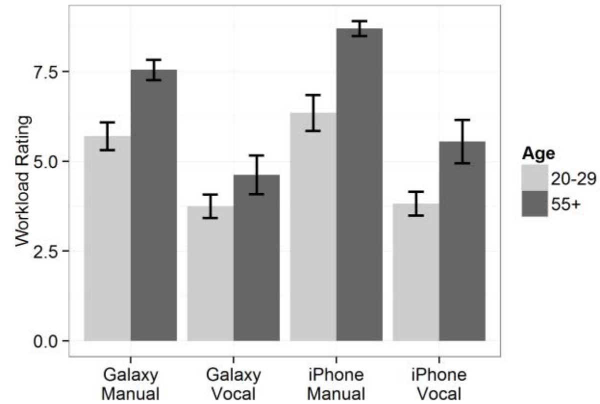

Figure 2. Workload ratings. Black bars represent SEM.

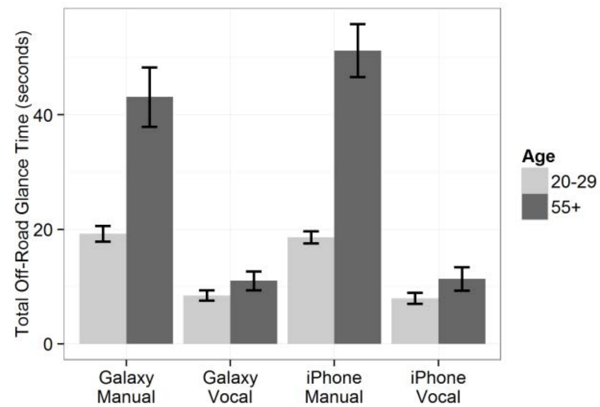

Figure 3. Off-road glance time. Black bars represent SEM.

The percentage change in vehicle velocity from single task driving shows a significantly larger decrease in speed for manual tasks compared to vocal (F(1,39)=5.44, p=.02). There was a higher standard deviation of lane position (SDLP) during manual input than vocal (F(1,39)=71.89, p<.001), as well as a significantly smaller SDLP for younger participants (F(1,38)=8.04, p=.007), but there was no significant age difference in percentage change in velocity. Perhaps most notable, participants demonstrated a significantly smaller SDLP while using the smartphone with the same OS as their primary phone (F(1,36)=5.65, p=.02) (Table 2). Participants demonstrated a nominally larger decrease in velocity when using their primary phone's OS.

DISCUSSION

Vocal input methods across the Galaxy and the iPhone show shorter task times, reduced off road glance time, and less presumed detrimental impact on vehicle control measures than manual interactions. These results are consistent with previous studies of voice interaction [3, 4, 8, 10], overall suggesting that vocal input can be less demanding than traditional visual-manual interactions for equivalent tasks. While nominal data values may suggest that one device might have a slight benefit over the other (e.g. shorter task time, etc.), these results are not statistically evident across the range of measures considered.

Reinforcing difficulties observed in the number of older participants that were withdrawn from the study, age effects were observed across a range of measures. Results suggest that older adults have more difficulty with both visual-manual and auditory-vocal tasks. The degree to which this is based upon prior experience is not clear in these findings. Additionally, typeface legibility research suggests that glance based reading is influenced by font style and age [6]; this may be one reason for the large age effect seen during visual-manual input, and the small differences seen between devices.

The influence of device familiarity is among the more notable results. Not surprisingly, due to previous experience, iPhone users perform better while using an iPhone and Android users perform better when using Android OS. The significant differences are especially remarkable considering this relatively small population. There is a clear trend suggesting that experience with a phone type (illustrated by phone ownership) plays a role in the overt demand observed in experimental assessments of this type.

CONCLUSION

Consistent with earlier work, this study illustrates the benefits of voice interfaces over manual interfaces while driving, while findings by device (iPhone vs Android) were suggestive of relatively similar demand profiles. The most striking finding is the suggestion of a general experience effect. This is supported by observations related to age as well as personal phone ownership (as a proxy for device familiarity). In summary, benefits of one device over another may be less grounded in the technology itself, but instead related to one's previous experience. Future work needs to further investigate assessed interface demands in naive and experienced user groups.

LIMITATIONS

This work excluded individuals with difficulty completing tasks and could be enhanced with a larger sample size. The study did not account for either device's ability to learn a user's speech patterns. This functionality is known to exist in Apple's Siri [1]. Participants in this data set own a variety of different phone models that can be broadly classified as iPhone or Android, but specific experience with the iPhone 5s or Galaxy S4 is unknown. No data were collected on the frequency of device use while driving.

ACKNOWLEDGEMENTS

Support for this work was provided by the New England University Transportation Center and the Toyota Class Action Settlement Safety Research and Education Program. The views and conclusions being expressed are those of the authors, and have not been sponsored, approved, or endorsed by Toyota or plaintiffs' class counsel.

REFERENCES

1. Apple Inc. (2015). About Siri - Apple Support. Retrieved from https://support.apple.com/en-us/HT204389
2. Dopart, C., Haggman, A., Thornberry, C., Mehler, B., Dobres, J., & Reimer, B. (2013). A Driving Simulation Study Examining Destination Entry with iOS 5 Google Maps and a Garmin Portable GPS System. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *57*(1), 1889–1893.
3. Lo, V. E.-W., & Green, P. A. (2013). Development and Evaluation of Automotive Speech Interfaces: Useful Information from the Human Factors and the Related Literature. *International Journal of Vehicular Technology*, *2013*, 1–13.
4. Maciej, J., & Vollrath, M. (2009). Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accident Analysis & Prevention*, *41*(5), 924–930.
5. Reimer, B. and Mehler, B. 2011. The impact of cognitive workload on physiological arousal in young adult drivers: a field study and simulation validation. *Ergonomics* 54, 10 (Oct. 2011), 932-942.
6. Reimer, B., Mehler, B., Dobres, J., Coughlin, J. F., Matteson, S., Gould, D., … Levantovsky, V. (2014). Assessing the impact of typeface design in a text-rich automotive user interface. *Ergonomics*, *57*(11), 1643–1658.
7. 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.
8. Reimer, B., Mehler, B., Reagan, I., Kidd, D. & Dobres, J. (2015). Multi-modal demands of a smartphone used to place calls and enter addresses during highway driving relative to two embedded systems. Insurance Institute for Highway Safety, Arlington, VA.
9. Rumschlag, G., Palumbo, T., Martin, A., Head, D., George, R., & Commissaris, R. L. (2015). The effects of texting on driving performance in a driving simulator: The influence of driver age. *Accident Analysis & Prevention*, *74*, 145–149.
10. Tsimhoni, O., Smith, D., & Green, P. (2004). Address entry while driving: speech recognition versus a touch-screen keyboard. *Human Factors*, *46*(4), 600–610.
11. Zhao, N., Reimer, B., Mehler, B., D'Ambrosio, L. A., & Coughlin, J. F. (2013). Self-reported and observed risky driving behaviors among frequent and infrequent cell phone users. *Accident Analysis & Prevention*, *61*, 71–77.