

# Comparison of Surface Gestures and Air Gestures for In-Vehicle Menu Navigation

Shawn Wu<sup>1</sup>, Thomas Gable<sup>2</sup>, Keenan May<sup>3</sup>, Young Mi Choi<sup>4\*</sup>, Bruce N. Walker<sup>5</sup>

<sup>1,2,3,5</sup> School of Psychology, Georgia Institute of Technology, Atlanta, USA

<sup>4</sup> School of Industrial Design, Georgia Institute of Technology, Atlanta, USA

<sup>5</sup> School of Interactive Computing, Georgia Institute of Technology, Atlanta, USA

---

## Abstract

**Background** Traditionally, in-vehicle systems have been operated using physical controls such as buttons and knobs. The array of tasks that these controls can efficiently accomplish in modern in-vehicle systems has outpaced their capabilities. The use of gestures for navigating interfaces has become increasingly prevalent solution to these shortcomings. However, best practices for the use of gestures have not been developed, leaving designers at risk of implementing systems that can increase a driver's workload and raise the risk of accidents.

**Methods** The current paper examines four gesture interaction techniques for completing a menu navigation task in terms of their effects on drivers' primary driving and secondary menu navigation performance, as well as self-reported workload and preferences. Participants used either a tap (point) gesture or a swipe gesture either in the air or on a gesture surface.

**Results** It was found that air gestures were slower than surface gestures and led to higher workload. Swipe gestures were lower workload than tap gestures, and the surface swipe gesture led to the lowest workload overall.

**Conclusions** These results indicate that, at present, surface gestures are a lower workload alternative to air gestures that retain some of their flexibility and naturalistic potential.

**Keywords** Gestures, In-Vehicle Interaction, Menu Navigation

---

\*Corresponding author: Young Mi Choi (christina.choi@gatech.edu)

*Citation:* Wu, S., Gable, T., May, K., Choi, Y., & Walker, B. N. (2016). Comparison of Surface Gestures and Air Gestures for In-Vehicle Menu Navigation. *Archives of Design Research*, 29(4), 65-81.

<http://dx.doi.org/10.15187/adr.2016.11.29.4.65>

**Received :** Aug. 11. 2016 ; **Reviewed :** Oct. 18. 2016 ; **Accepted :** Oct. 18. 2016

**pISSN** 1226-8046 **eISSN** 2288-2987

**Copyright :** This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted educational and non-commercial use, provided the original work is properly cited.

---

## 1. Introduction

Vehicles are becoming as much spaces for mobile computing as modes of transportation. While drivers have been placing phone calls and interacting with navigation systems for years, they are now able to interact with complex multifunction infotainment systems such as CarPlay or Android Auto [1] that reproduce the varied capabilities of modern smart phones within the in-vehicle interface. While these platforms do reflect efforts to adapt mobile or desktop paradigms to the in-vehicle context through modifications such as larger icons, fundamentally, those paradigms were not developed explicitly to support multi-tasking. In general, humans have difficulty carrying out multiple simultaneous tasks without degrading performance [18]. As such, performing the secondary task of interacting with in-vehicle menu systems while driving can lead to accidents. Dingus et al. [8] found that almost 80% of car crashes involved a distraction from driving in the three seconds immediately prior to the accident, and interaction with in-vehicle computing devices being the most common mode of distraction.

In-vehicle interfaces must adapt in a variety of ways to increase the ease with which a driver might utilize them while driving. These include structuring tasks to allow flexible interleaving [25], minimizing visual demand [9,19], minimizing demand associated with manual execution such as making precise motor movements [19], and distributing driver workload across sensory modalities [19]. Relating to all of these efforts is the need to lower the cognitive workload incurred by drivers who are using secondary computing interfaces. High cognitive load has been found to be predictive of many measures of dual-task performance in the vehicle [4, 24, 9].

One avenue that researchers have explored for optimized in-vehicle interfaces in the aforementioned ways is the study of different of input controls. Traditionally, in-vehicle systems have been operated using physical controls such as buttons and knobs. These controls have certain benefits, including the rich tactile feedback they can afford, the ability of drivers to learn the locations of different buttons, and the fact that drivers may already be familiar with using them. However, the array of computing tasks that can be efficiently accomplished using these controls has outpaced their capabilities.

Direct (aimed) touch control systems, in which users make commands by pressing virtual buttons on a screen, are currently the most common solution to the need to support complex, multifunction interfaces in the vehicle. While these interfaces may be highly learnable and conceptually simple, they have been found to be more visually demanding than traditional buttons and knobs, and as such have been shown to lead decreased driving performance [5, 26]. Noy et al. [21] found that the lack of tactile feedback with touch screens often forces drivers to use their vision to locate controls whose locations have not been memorized.

Speech input controls have also become more common. However, while they may alleviate visual demand on drivers, speech interfaces have been found to lead to large increases in cognitive workload [9], they may not function well in noisy conditions, may not function well where internet connectivity is poor, and are inappropriate for many tasks such as information browsing or parameter adjustment.

In response to these shortcomings with direct touch and voice controls, interfaces that leverage the communicative power and flexibility of surface or air gestures are becoming common [22]. Gestures have a certain basic appeal- Akyol et al. [2] found that participants conceptualize gesture interactions as a natural and intuitive interaction method. Carrying out gestures does not require the visually demanding act of aimed pointing of direct touch systems. Additionally, gestures may not be affected by the same recognition, cognitive workload, or task-appropriateness issues that limit the utility of current speech interfaces. Finally, air gestures have the advantage of not being requiring that the user locate a screen or specific physical control [23]. Thus, they are being explored as a potential alternative for drivers to carry out secondary menu navigation. Automobile manufacturers such as Audi, BMW, Ford and Mercedes-Benz are working on developing air gesture control systems for their vehicles [7]. However, best practices for the use of surface and air gestures have not been developed. This leaves designers at risk of relying on their apparent intuitiveness and creating systems that, in fact, raise driver workload and accident risk.

### **1. 1. Surface Gestures**

A variety of research has been done on surface gestures in the vehicle, carried out either on a head unit screen or on an external gesture surface. Christiansen et al. [6] compared two swipe-based gestural menu input systems with two direct touch systems, and found that the gestural systems led to significantly fewer glances away from the road. More tellingly, the gesture systems led to far fewer 'long' glances between 0.5 and 2 seconds. The authors explained this difference by the fact that direct touch systems tend to require extended glances to locate a target and, subsequently, to check the location of the hand relative to the target. Gestures, on the other hand, require only brief glances to locate the input surface (in this case the head unit screen). Similarly, Jaeger et al. [13] compared traditional tactile buttons, direct touch, and a swipe-based surface gesture interface. They found that the traditional tactile interfaces led to greater lane exiting errors and steering wheel turning, and caused the greatest number of eye glances away from the forward roadway. The gesture-based interface required the least amount of visual load (least number of eye glances), while the direct touch interface was the most time efficient for completing secondary tasks. Swette, May, Gable, and Walker [25] compared several surface gesture interfaces to a direct touch system and found that a swipe-based system led to lowered visual distraction compared to other gesture systems and the direct touch system. This finding was mirrored by Kujala [15], who compared several methods of surface gesture scrolling: kinetic gestural scrolling, discrete page-by-page swipe scrolling, and direct touch up and down arrows. It was again found that the discrete swipe gesture system was highest performing. Thus, it may be concluded that swipe-based surface gestures are a viable option for in-vehicle interface design.

## 1. 2. Air Gestures

Research into the efficacy of air gestures has generally been less concrete. Multiple studies have found that users tended to prefer simple directional air gesture interfaces to tactile buttons, since this required less visual input and users didn't have to touch or reach anything [3, 9]. May, Gable, and Walker [16] implemented a swipe-based air gesture system and compared this to a direct touch system. They found that the air gesture system led to less visual distraction but higher self-reported cognitive load and task times compared with the direct touch system. They noted that air gestures run the risk of creating high workload for drivers as they try to recall and precisely execute required gestures [10]. As a whole, prior work on air gestures indicates that while they may have advantages over current touch systems in terms of the amount of visual distraction they pose and the extent to which users enjoy their use, cognitive load is an ever-present concern.

## 1. 3. Guidelines for Surface and Gestures

Some guidelines do exist for in-vehicle gestures. Gable, May and Walker [10] recommended that gesture systems use limited sets of relative-mapped gestures to avoid the high working memory load for recall and execution of a large set of absolute-mapped gestures. Such systems, in which users repeatedly swipe or point to move through menus in a serial fashion, have been evaluated both in the context of air gestures and in the context of surface gestures. These systems can be designed to sense either surface gestures on a screen or touchpad (2D) or air gestures (3D). While both gesture mediums can be used to support similar sets of basic gestures, each medium may impose different levels of cognitive demand. Thus far, little research has addressed these differences.

## 1. 4. Current Study

While previous research has focused on comparing gestures with other modalities of input and attempting to assess the viability of surface and air gestures in general, few studies have addressed the difference between air gestures and surface gestures, particularly across types of gestures likely to be used by designers. May, Gable, and Walker [17] conducted a two-part participatory design and workload assessment activity on various air gestures for menu navigation in the vehicle. They found that self-reported workload for incremental movement tasks was lowest using simple swipe gestures. In that study, pointing gestures were also commonly suggested by users and found to be low workload, suggesting that pointing and swiping are two viable alternatives for simple symbolic relative-mapped gesture input for menu browsing.

The current study focuses on these two types of gestures (swiping and pointing), as instantiated in air gesture and surface gesture systems. Performance of gestural menu control systems that were actuated either through swipe or tap (point) gestures, carried out either on a touch surface or in the air were evaluated. Performance was evaluated through primary driving performance, secondary menu task performance, self-reported cognitive workload, and preferences.

---

## 2. Methods

### 2. 1. Participants

A total of 14 participants with a mean age of 24.81 (SD= 1.5) took part in the study. The inclusion criteria required all participants to have had a valid drivers' licenses for two years or more, and have normal or corrected to normal vision and hearing. None of the 14 participants had experience with driving simulators prior to the study, and 3 of the participants had used a Leap Motion device prior to the study.



Figure 1 The driving simulator and Leap Motion sensor with secondary display

### 2. 2. Apparatus

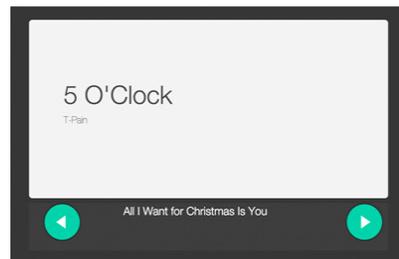
#### 2. 2. 1. Driving Simulator

The Simulator was performed using low fidelity simulator OpenDS. Simulator visuals were displayed on a 40" LED monitor and input was given using a Logitech Driving Force GT racing wheel and floor pedals (Figure 1). The Three-Vehicle-Platoon task was used for the primary driving task. Participants were instructed to maintain central lane position (maintaining latitudinal control), while following a lead car at a 20-meter distance (maintaining longitudinal control). The lead car would periodically brake and lower its speed. Participants were instructed to use their brakes and strive to maintain following distance in respond to these events. The values of longitudinal and lateral deviation, as well as brake reaction time were all automatically recorded and stored.

#### 2. 2. 2. Secondary Menu Interface

Four types of interactions were used in the study: air swipe (AS), air tap (AT), touch screen swipe (SS), and touch screen tap (ST). The same visual and audio cues were given to users for the four interaction types in order to provide feedback regarding the current soundtrack and their task. The graphical interface (Figure 2) had a green "previous" button on the left of the screen and another on the right for "next," while the currently selected song was displayed in the center. The current goal song was displayed at the bottom of the interface in case the participants forgot what they were looking for. As pointed out by Gable, May and Walker [9], audio cues within gesture interfaces are likely to make them more usable and decrease visual demand, so they were included in all gesture conditions in the current study.

The audio cues used in the system gave the user information of the currently selected song using spindex [13] and text to speech (TTS) audio cues.



**Figure 2** The current song displayed in large type, and the target song at the bottom of the screen in small type

To display this information an iPad Mini was used as the screen, and stereo speakers for auditory output. For input in the air gesture systems, a Leap Motion device carried out the recognition process. The Leap is a small device that utilizes infrared light and depth cameras for fingertip-precision gesture recognition within about 3x3x3 foot volume. As seen in Figure 1, the device was placed on the right side of participant, in front of the secondary display. For the two surface touch interfaces, the multi-touch recognition of the iPad Mini was employed. During the air gesture conditions, the interaction with the Leap and the user interface were powered by a MacBook Air whose visuals were relayed via an application called “air display” to the iPad Mini.

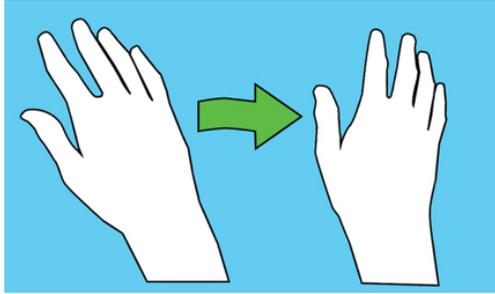
The secondary task simulator was a Javascript web application. The program held a 50-item library of audio tracks sorted and displayed alphabetically. The difficulty of each music selection task ranged from level 1 to 20, whereas level 1 indicated finding the first song in the library. Task times were recorded by the software, with the eight second selection time removed.

## 2. 3. Gesture Types

In order to move through the menus, participants used one of four interaction methods.

### 2. 3. 1. Air Swipe

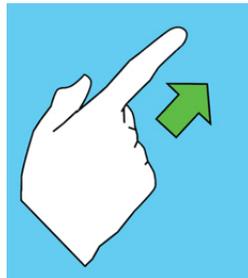
For the air swipe (AS) gesture, swiping to the right in the air meant 'next' while swiping towards the left meant 'previous' (Figure 3). Using the taxonomies of May, Gable and Walker [17], AS was a dynamic discrete gesture whose motion was defined by the speed, magnitude and direction of hand movement. The metaphor was direct manipulation, and the spatial mapping related the start of the menu with the left of the user and the end of the menu with their right. Many repeated actions were required to complete the task, and no hold state was used. Participants were not required to take any specific pose with their hand. Finally, this gesture involved intentionally controlling the wrist+arm but not individual digits.



**Figure 3** The 'air swipe' gesture: carried out by rotating the wrist or pivoting at the elbow or wrist (hand angle/ pose was not sensed). A right swipe moved to the next song, while a left swipe moved to the previous song

### 2. 3. 2. Air Tap

To actuate an air tap (AT), participants used their fingertips to tap forward in the air (Figure 4). The lateral direction of their forefingers determined the action of the secondary task, with tapping to the right meaning 'next' and tapping to the left 'previous'. This was a dynamic discrete gesture that used hand pose, direction of movement, as well as magnitude and speed of movement as codes for the gesture. The metaphor was either deictic signaling or mimicking a physical device (such as the surface tap interface's buttons). The spatial mapping related the start of the menu with the left of the user and the end of the menu with their right. Many repeated actions were required to complete the task, and no hold state was used. Finally, this gesture required controlling the wrist, arm and digits.



**Figure 4** The 'air tap' gesture. Pointing right moved to the next song, and pointing left moved to the previous song

### 2. 3. 3. Surface Gestures

The same gestures were used for the surface swipe (SS) and surface tap (ST) interactions, but instead of participants performing the gestures in the air participants simply swiped anywhere on the touchscreen for SS or touched buttons at the bottom left or bottom right of the screen for ST (Figure 2).

### 2. 3. 4. Making a Selection

To select an item, participants simply ceased any interaction with the interface and waited eight seconds, after which the system would choose the currently shown song. This method of selection is similar to many traditional multimedia browsing interfaces in which the driver does not need to press the 'play' button every time they change the soundtrack, and instead press 'next' or 'previous' until they see/hear a song they want. Eight seconds was chosen for the duration needed for selection since it was long enough for participants to put their hands back on the wheel and make any corrections during their driving task and then switch back to the selection task if needed.

## **2. 4. Survey Instruments**

### **2. 4. 1. Subjective Workload**

In order to measure participants' perceived workload for each interaction type, the NASA TLX survey [12] was employed. The survey asked participants to rate a task on six axes: mental demand, physical demand, temporal demand, performance, effort, and frustration. These were then combined to produce an unweighted composite 'subjective workload' score out of 100, with higher scores indicating higher perceived workload. Additionally, each component was analyzed individually.

### **2. 4. 2. Preferences**

An additional questionnaire was used to measure participants' subjective preferences between the interfaces. Participants were asked to rank the four interfaces based on their perceived overall preference, effectiveness, efficiency, and satisfaction.

## **2. 5. Procedure**

When participants first arrived to the study they completed an informed consent procedure and went through the instructions for the study with the experimenter. They were then given a verbal and visual introduction to the four interaction types (AS, AT, SS, and ST). Participants were then given a tutorial about the driving simulator and the task they were to complete with it. After the participants stated to the experimenter that they were comfortable with the driving simulator, they were given a more in-depth walkthrough of the secondary task and then began the experimental conditions.

A total of five conditions were performed in the study, including a driving-only rest drive and a dual-task driving + menu navigation task for each of the four interfaces. In order to distribute variance associated with order effects, the sequence of the five conditions was counterbalanced using a Latin Square method. Before the experiment began, all the participants were informed that they should prioritize the driving task but attempt to complete the secondary task as quickly and accurately as possible.

Prior to each experimental condition, participants were given a short practice session (8 search tasks). In both the practice and the experimental conditions the first task started after the participants' car passed the starting sign in the simulated environment. There were 12 search tasks for each condition. A break of 3-5 seconds was given between each search task. After each condition was completed, the participants completed the NASA TLX survey. After participants completed all five conditions, they were asked to complete the personal preferences and demographic questionnaire.

## **2. 6. Analyses**

For each dependent variable, a 2x2 Huynh-Feldt repeated measures Analyses of Variance was conducted. The two factors were gesture medium, with levels 'air' and 'surface', and gesture motion, with levels 'swipe' and 'tap.' Preferences were analyzed using Wilcoxon signed-rank tests (the nonparametric equivalent of a paired samples two-tailed t-test).

## 2. 7. Hypotheses

It was hypothesized that air gestures would be more preferred, but would be more time consuming and lead to higher workload, compare to surface gestures, and that swipe gestures would overall be better performing than tap gestures.

---

## 3. Results

### 3. 1. Subjective Workload

#### 3. 1. 1. Composite Workload

There was a significant main effect of gesture medium,  $F(1,13) = 22.999$ ,  $p < .001$ ,  $\eta^2 = .639$ . Air gestures, ( $M = 59.33$ ,  $SD = 15.26$ ), led to higher composite workload compared to surface gestures, ( $M = 36.21$ ,  $SD = 19.27$ ) (Figure 5). There was also a significant main effect of gesture motion,  $F(1,13) = 11.211$ ,  $p = .005$ ,  $\eta^2 = .463$ . Tap gestures, ( $M = 53.27$ ,  $SD = 22.74$ ), led to higher composite subjective workload than swipe gestures, ( $M = 42.27$ ,  $SD = 17.36$ ). The interaction of Medium and Motion was not significant,  $F(1,13) = 2.238$ ,  $p = .159$ .

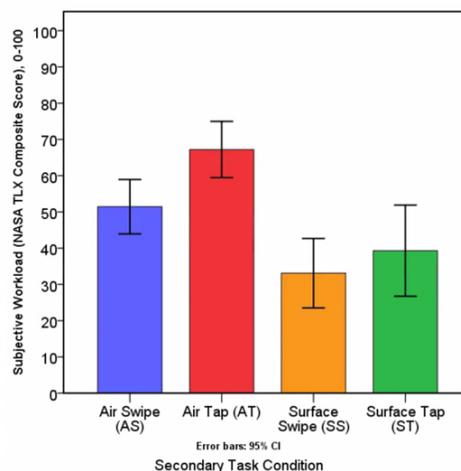


Figure 5 Subjective Workload (NASA TLX Composite Score, 0–100, smaller is better) by secondary task condition

#### 3. 1. 2. Component Workload Subscales

A Hyunh-Feldt two-way repeated measures ANOVA was conducted on each component of the NASA TLX. The ‘performance’ component was not analyzed because objective performance measures were present.

For the NASA TLX mental component, there was a significant main effect of gesture medium,  $F(1,15) = 9.244$ ,  $p = .008$ ,  $\eta^2 = .381$ . Air gestures ( $M = 59.66$ ,  $SD = 13.32$ ), led to a higher self-reported mental workload compared to surface gestures ( $M = 39.84$ ,  $SD = 18.68$ ). There was not a significant main effect of gesture motion,  $F(1,15) = 2.358$ ,  $p = .145$ . The interaction of Medium and Motion was also not significant,  $F(1,15) = 0.792$ ,  $p = .388$ .

For the NASA TLX physical component, there was a significant main effect of gesture medium,  $F(1,15) = 25.230$ ,  $p < .001$ ,  $\eta^2 = .627$ . Air gestures ( $M = 65.88$ ,  $SD = 15.24$ ) led to a higher self-reported physical workload compared to surface gestures ( $M = 36.34$ ,  $SD = 19.96$ ). There was not a significant main effect of gesture motion,  $F(1,15) = 3.945$ ,  $p = .066$ , although this appeared to be trending toward significance, and observed power was relatively small (.460). With more participants, it is likely that tap gestures ( $M = 54.09$ ,  $SD = 15.24$ ) may have led to higher self-reported physical workload compared to swipe gestures ( $M = 48.13$ ,  $SD = 13.92$ ). Finally, the interaction of Medium and Motion was not significant,  $F(1,15) = 1.029$ ,  $p = .289$ .

For the NASA TLX temporal component, there was a significant main effect of gesture medium,  $F(1,15) = 7.151$ ,  $p = .017$ ,  $\eta^2 = .323$ . Air gestures ( $M = 58.25$ ,  $SD = 13.48$ ) led to a higher self-reported temporal workload compared to surface gestures ( $M = 40.97$ ,  $SE = 20.76$ ). There was also a significant main effect of gesture motion,  $F(1,15) = 5.334$ ,  $p = .036$ ,  $\eta^2 = .323$ . Tap gestures ( $M = 54.84$ ,  $SD = 11.60$ ) led to higher self-reported temporal workload compared to swipe gestures, ( $M = 44.38$ ,  $SD = 15.28$ ). However, the interaction of Medium and Motion was not significant,  $F(1,15) = 0.147$ ,  $p = .708$ .

For the NASA TLX effort component, there was a significant main effect of gesture medium,  $F(1,15) = 14.048$ ,  $p = .002$ ,  $\eta^2 = .484$ . Air gestures, ( $M = 62.13$ ,  $SD = 11.60$ ), led to a higher self-reported effort compared to surface gestures ( $M = 42.84$ ,  $SD = 26.48$ ). There was also a significant main effect of gesture motion,  $F(1,15) = 5.151$ ,  $p = .038$ ,  $\eta^2 = .256$ . Tap gestures ( $M = 57.56$ ,  $SD = 12.84$ ) led to higher self-reported effort compared to swipe gestures ( $M = 47.01$ ,  $SD = 15.28$ ). However, the interaction of Medium and Motion was not significant,  $F(1,15) = 1.775$ ,  $p = .203$ .

For the NASA TLX frustration component, there was a significant main effect of gesture medium,  $F(1,15) = 13.826$ ,  $p = .002$ ,  $\eta^2 = .480$ . Air gestures ( $M = 50.56$ ,  $SD = 13.40$ ) led to a higher self-reported frustration compared to surface gestures ( $M = 30.81$ ,  $SD = 18.80$ ). There was also a significant main effect of gesture motion,  $F(1,15) = 10.695$ ,  $p = .005$ ,  $\eta^2 = .416$ . Tap gestures, ( $M = 47.25$ ,  $SD = 16.40$ ) led to higher self-reported frustration compared to tap gestures, ( $M = 34.13$ ,  $SD = 13.40$ ). However, the interaction of Medium and Motion was not significant,  $F(1,15) = 2.855$ ,  $p = .112$ .

### 3. 2. Primary Driving Task Performance

Participants' longitudinal and lateral deviation as well as their brake response time were measured by the driving simulator and averaged across each condition (see Table 1).

Table 1 Driving performance (longitudinal deviation, lateral deviation, and brake response time. Smaller is better.)

Condition	Air		Surface		Mean	
	M	(SD)	M	(SD)	M	(SD)
<b>Longitudinal dev.</b>						
Swipe	22.61	(2.51)	22.23	(2.28)	22.42	(2.36)
Tap	22.58	(2.91)	23.00	(3.03)	22.79	(2.92)
Mean (SD)	22.60	(2.66)	22.62	(2.66)	-	-
<b>Lateral deviation</b>						
Swipe	0.497	(0.253)	0.448	(0.231)	0.472	(0.239)
Tap	0.504	(0.177)	0.513	(0.249)	0.509	(0.212)
Mean (SD)	0.500	(0.215)	0.481	(0.238)	-	-
<b>Brake RT (ms)</b>						
Swipe	1264	(111)	1239	(156)	1252	(133)
Tap	1293	(140)	1322	(177)	1307	(158)
Mean (SD)	1278	(125)	1280	(169)	-	-

### 3. 2. 1. Longitudinal Deviation

There were no significant main effects of gesture medium,  $F(1,13) = .003$ ,  $p = .960$ , or gesture motion on longitudinal deviation (following distance),  $F(1,13) = .483$ ,  $p = .499$ . The interaction of Medium and Motion was also not significant,  $F(1,13) = .735$ ,  $p = .407$ .

### 3. 2. 2. Lateral Deviation

There were no significant main effects of gesture medium,  $F(1,13) = .131$ ,  $p = .724$  or gesture motion,  $F(1,13) = 2.563$ ,  $p = .133$  on lateral deviation (lane keeping). The interaction of Medium and Motion also not significant,  $F(1,13) = .973$ ,  $p = .342$ .

### 3. 2. 3. Brake Response Time

There were no significant main effect of gesture medium,  $F(1,13) = .004$ ,  $p = .948$ , or of gesture motion,  $F(1,13) = 4.414$ ,  $p = .056$ , on brake response time. However, the data appeared to trend toward significance, and observed power was very low (.050). It is thus likely that with a larger sample size, swipe gestures ( $M = 1.25s$ ,  $SD = 0.13$ ), would have been found to lead to lower brake response time than tap gestures, ( $M = 1.31s$ ,  $SD = 0.16$ ).

## 3. 3. Secondary Task Performance: Task Completion Time

There was a significant main effect of gesture medium on task completion time,  $F(1,13) = 38.31$ ,  $p < .001$ ,  $\eta^2 = .747$ ). Using air gestures ( $M = 19.29s$ ,  $SD = 7.25$ ) led to longer task times compared to surface gestures ( $M = 11.12s$ ,  $SD = 2.10$ ). There was not a significant main effect of gesture motion,  $F(1,13) = 2.12$ ,  $p = .169$ , nor was the interaction of Medium and Motion significant,  $F(1,13) = 1.51$ ,  $p = .104$ .

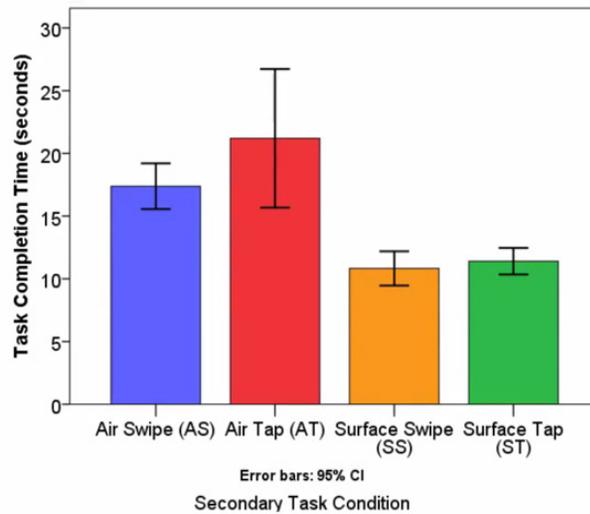


Figure 6 Average Task Completion Time (in seconds, smaller is better) by secondary task condition

### 3. 4. Preference Ratings

#### 3. 4. 1. Overall Preference

The summary of the subjective ratings can be seen in Table 2. The analyses found that for overall preference, SS (Md = 1) was most preferred. It was ranked significantly higher than AT (Md = 4),  $z = -3.508$ ,  $p < .001$ , AS (Md = 2.5),  $z = -3.040$ ,  $p = .02$ , and ST (Md = 2.5),  $z = -2.037$ ,  $p = .042$ .

#### 3. 4. 2. Effectiveness

In terms of effectiveness, SS (Md = 1) was most preferred. It was ranked significantly higher than AS (Md = 3),  $z = -3.01$ ,  $p = .003$  and AT (Md = 4),  $z = -3.46$ ,  $p = .001$ .

#### 3. 4. 3. Efficiency

In terms of efficiency, SS (Md = 1) was most preferred. It was ranked significantly higher than AS (Md = 3),  $z = -3.22$ ,  $p = .001$  and AT (Md = 4),  $z = -3.34$ ,  $p = .001$ .

#### 3. 4. 4. Satisfaction

With regard to satisfaction, SS (Md = 1) was ranked again as most preferred. It was ranked higher than AS (Md = 2),  $z = -2.80$ ,  $p = .005$ , ST (Md = 2)  $z = -2.49$ ,  $p = .013$ , and AT (MD = 4),  $z = -3.44$ ,  $p = .001$ .

Table 2 Subjective rankings of gestures (smaller is better)

Median Rating	Air Swipe (AS)	Air Tap (AT)	Surface Swipe (SS)	Surface Tap (ST)
Overall preference	2.5	4	1	2.5
Effectiveness	3	4	1	2
Efficiency	3	4	1	2
Satisfaction	2	4	1	3

---

## 4. Discussion and Conclusion

### 4. 1. Relation of Results to Hypotheses

The current study examined ST, SS, AS, and AT in terms of primary driving performance, secondary task performance, subjective workload, and preferences. As hypothesized, surface gestures and swipe gestures were higher performing in terms of workload and completion time, and the intersection of these, surface-swipe, was the most preferred interaction method. In addition, surface gestures may have led to significantly shorter brake response times given more participants. However, contrary to expectations and prior findings, participants did not prefer the air gesture interfaces.

Drilling down into the observed differences in NASA TLX components, air gestures led to higher reported workload across many components. Of these, the difference between gesture mediums was the largest for physical workload and frustration. This indicates that a primary source of difficulty associated with air gestures was in their physical execution, and that some frustration was present, perhaps due to this difficulty. This frustration associated with physical execution difficulty would explain the lower preference rankings of the air gesture interfaces.

For gesture motion, swipe gestures were better-performing than tap gestures across several NASA TLX components, with the largest effects being found for frustration, followed by temporal load. This pattern is more difficult to interpret, but could indicate that participants had difficulty executing the tap gestures consistently, which increased the perceived difficulty of interleaving primary and secondary tasks.

### 4. 2. Limitations

The poor performance of air gestures in this study may have been due in part to the air gesture system having trouble recognizing gestures, rather than difficulty inherent to using these gestures. However, these two sources of difficulty are not distinct from one another. A 'perfect' gesture system could read a person's intent so competently that there would be little difficulty associated with physical execution. When and if such a system is created, workload will stem more from the difficulty of recalling gestures, getting feedback about the effects of each gesture, and whether gestural interaction styles allow drivers to interleave tasks in a way that minimizes performance costs. While these issues are still germane to present gesture designers, the present results suggest that recognition technology has not progressed to a point where physical execution can be carried out effectively without the full attention of the user. The sensor technology used in this study was state of the art. While participants had trouble consistently triggering desired air gestures, they were successful often enough that task completion times with the air gesture systems were comparable (although admittedly longer) to those required for the surface gesture systems. However, the trending difference in brake response times indicates that performance in the driving task may have degraded due to the increase in workload required to achieve even this level of performance.

Thus, it appears that it is one thing to construct a functioning air gesture system, and another to create such a system that functions well even in the face of user inattention and distraction. Taken as whole, these results indicate that air gestures are, at present, prohibitively high-workload for use by the multi-tasking driver. Future improvements to sensors and software might produce a different pattern of results. In the short term, these findings speak to the need to define gestures in a flexible manner that allows natural variation in the physical execution of each gesture to be accommodated, and, accordingly, to the need for smaller gesture sets that allow for such flexible definitions [16]. Over time, issues such as memorability and visibility of feedback may become more important than optimizing physical execution difficulty to accommodate the limitations of current sensors.

It is important to note that the sample size for this study was small, and statistical power was low in many cases. Additionally, while it has utility as a readily accessible research tool, the OpenDS simulator is not high fidelity and the Three-Vehicle Platoon driving task does not reflect the complexity of a real driving environment. Future work should endeavor to evaluate whether these findings can be extended to a more realistic driving environment, with a larger sample. From a methodological standpoint, design researchers may note that a variety of useful results were observed using the NASA TLX survey, despite the relatively small sample size. Since cognitive workload underpins many other measures of safe driving, this tool presents a viable option when resources are limited.

### 4. 3. Implications

The present work suggests that in-vehicle interface designers should eschew aimed pointing gestures, and instead focus on integrating easily triggered dynamic-discrete gestures to support in-vehicle information browsing. Air gestures should be used with caution until sensor technology improves to the point where physical execution does not lead to elevated workload and frustration. If air gestures are used, maximizing the ability of users to physically execute them without errors may be more important than optimizing other features of the interaction. Surface gestures offer similar flexibility to accommodate the range of input needs of the modern driver, but do not suffer as badly from issues with physical execution and overall workload. Thus, surface gestures are recommended as a lower-workload alternative for future interactions in the vehicle.

### References

- 1 Amadeo, R. (2016, January 24). CarPlay vs Android Auto: Different approaches, same goal. Retrieved October 05, 2016, from <http://arstechnica.com/gadgets/2016/01/carplay-vs-android-auto-different-approaches-same-goal/>.
- 2 Akyol, S., Canzler, U., Bengler, K., & Hahn, W. (2000, November). Gesture Control for Use in Automobiles. In *MVA* (pp. 349–352).
- 3 Alpern, M., & Minardo, K. (2003, April). Developing a car gesture interface for use as a secondary task. In *CHI'03 extended abstracts on Human factors in computing systems* (pp. 932–933). ACM.
- 4 Antin, J. F., Dingus, T. A., Hulse, M. C., & Wierwille, W. W. (1990). An evaluation of the effectiveness and efficiency of an automobile moving-map navigational display. *International Journal of Man-Machine Studies*, 33(5), 581–594.

- 5 Bellotti, F., De Gloria, A., Montanari, R., Dosio, N., & Morreale, D. (2005). COMUNICAR: designing a multimedia, context-aware human-machine interface for cars. *Cognition, Technology & Work*, 7(1), 36-45.
- 6 Christiansen, L. H., Frederiksen, N. Y., Jensen, B. S., Ranch, A., Skov, M. B., & Thiruravichandran, N. (2011). Don't look at me, I'm talking to you: investigating input and output modalities for in-vehicle systems. In *Human-Computer Interaction-INTERACT 2011* (pp. 675-691). Springer Berlin Heidelberg.
- 7 Damon Lavrinc. (2014, July). Gesture Controls Are Coming To New Cars Next Year. From: <http://jalopnik.com/gesture-controls-are-coming-to-new-cars-next-year-1598692093>.
- 8 Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J. D., & Bucher, C. (2006). The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment (No. HS-810 593).
- 9 Engström, J., Johansson, E., & Östlund, J. (2005). Effects of visual and cognitive load in real and simulated motorway driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(2), 97-120.
- 10 Gable, T. M., May, K. R., & Walker, B. N. (2014, September). Applying Popular Usability Heuristics to Gesture Interaction in the Vehicle. In *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1-7). ACM.
- 11 Greenberg, J., Tijerina, L., Curry, R., Artz, B., Cathey, L., Kochhar, D., ... & Grant, P. (2003). Driver distraction: Evaluation with event detection paradigm. *Transportation Research Record: Journal of the Transportation Research Board*, (1843), 1-9.
- 12 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.
- 13 Jæger, M. G., Skov, M. B., & Thomassen, N. G. (2008, April). You can touch, but you can't look: interacting with in-vehicle systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1139-1148). ACM.
- 14 Jeon, M., Walker, B. N., & Srivastava, A. (2012). "Spindex" (Speech Index) Enhances Menus on Touch Screen Devices with Tapping, Wheeling, and Flicking. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 19(2), 14.
- 15 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.
- 16 May, K. R., Gable, T. M., & Walker, B. N. (2014, September). A multimodal air gesture interface for in vehicle menu navigation. In *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1-6). ACM.
- 17 May, K., Gable, T. M., & Walker B. N. (2016, under review). Recommended Air Gestures for Control of In-Vehicle Menus with Minimal Workload. *International Journal of Human-Computer Studies*.
- 18 Müller, C., & Weinberg, G. (2011). Multimodal input in the car, today and tomorrow. *IEEE Multi Media*, (1), 98-103.
- 19 National Highway Traffic Safety Administration. (2012). Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. Washington, DC: National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT).
- 20 Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J., & Goodman, M. (2005). An overview of the 100-car naturalistic study and findings. National Highway Traffic Safety Administration, Paper, (05-0400).
- 21 Noy, Y. I., Lemoine, T. L., Klachan, C., & Burns, P. C. (2004). Task interruptability and duration as measures of visual distraction. *Applied Ergonomics*, 35(3), 207-213.
- 22 Parada-Loira, F., González-Agulla, E., & Alba-Castro, J. L. (2014, June). Hand gestures to control infotainment equipment in cars. In *Intelligent Vehicles Symposium Proceedings, 2014 IEEE* (pp. 1-6). IEEE.

- 23 Pickering, C. A., Burnham, K. J., & Richardson, M. J. (2007, June). A research study of hand gesture recognition technologies and applications for human vehicle interaction. In *3rd Conf. on Automotive Electronics*.
- 24 Schaap, T. W., Van der Horst, A. R. A., Van Arem, B., & Brookhuis, K. A. (2013). *The relationship between driver distraction and mental workload* (Vol. 1, pp. 63–80). Farnham: Ashgate.
- 25 Swette, R., May, K. R., Gable, T. M., & Walker, B. N. (2013, October). Comparing three novel multimodal touch interfaces for infotainment menus. In *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 100–107). ACM.