Transportation Research Part F 51 (2017) 38–46

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

Using tactile detection response tasks to assess in-vehicle voice control interactions

Chun-Cheng Chang [a], Linda Ng Boyle [a,*], John D. Lee [b], James Jenness [c]

[a] University of Washington, Seattle, WA, United States
[b] University of Wisconsin-Madison, WI, United States
[c] Westat, Rockville, MD, United States

ARTICLE INFO

Article history:
Received 20 April 2016
Received in revised form 5 December 2016
Accepted 9 June 2017
Available online 30 September 2017

Keywords:
Cognitive distraction
Driving simulator
Voice control
Tactile detection response task

ABSTRACT

In-Vehicle Information Systems (IVIS) are becoming more accessible to drivers but also contain more complex communication features. Voice control systems are shown to be less distracting than visual-manual interfaces, but they can still impose cognitive workload. This study examined the cognitive workload associated with interactions with voice systems while driving. Cognitive workload was assessed using a Tactile Detection Response Task (TDRT) protocol. A driving simulator study with 48 participants was conducted using an interface with a Wizard-of-Oz based voice control system. Drivers conducted several voice tasks that included radio channel selection, address navigation, and scheduling a calendar appointment. Recognition accuracy and system delay time was manipulated within these voice tasks. Using a mixed linear model, cognitive workload was shown to be higher for navigation and calendar tasks when compared to radio tasks. Recognition errors and time delays in the calendar task significantly decreased TDRT response time. Drivers that are distracted by voice control systems that also contained system delays and errors do not necessarily increase TDRT response time. In fact, drivers may adapt over time to these system imperfections.

1. Introduction

In-vehicle voice control systems (VCS's) help reduce driver distractions as they allow drivers to operate their infotainment and other vehicle controls without taking their eyes off the road or their hands off the wheel (Putze & Schultz, 2012). The ease of VCS interactions make it possible for car manufacturers to offer drivers more complex features that would not be possible to perform while driving with a visual-manual interface. Such features may include address entry and message composition while driving. In practice, VCS may require button presses and glances toward the in-vehicle display. There are also issues associated with speech-recognition errors and response delays that may divert drivers attention away from the road where driver performance and safety is compromised.

Voice-based systems are shown to be less distracting than systems that require visual-manual interactions (Carter & Graham, 2000; Tijerina et al., 1998). However, they are not necessarily distraction free. Lee, Caven, Haake, and Brown (2001) found that drivers's reaction time to a lead vehicle braking event increased by 30% when using a voice control system compared to a baseline condition. Engström, Johansson, and Östlund (2005) found that participants who had to perform an

* Corresponding author.
E-mail address: linda@uw.edu (L.N. Boyle).

https://doi.org/10.1016/j.trf.2017.06.008
1369-8478/

auditory continuous memory task (n-back) had gazes that were concentrated more towards the center of the road, which limited their attention to other areas.

Visual and manual distractions might be reduced with the use of VCS, but these systems still impose cognitive workload on the driver. Cognitive workload is defined as the amount of attentional resources required to perform a task (De Waard, 1996). It is important to be able to quantify cognitive workload imposed by the interactions with an in-vehicle interface because it provides insights on the limitations of human attentional resources and how the attention is allocated (prioritized) by the humans in their management of the different tasks (Patten, Kircher, Östlund, & Nilsson, 2004).

The Detection Response Task (DRT) is a relatively low-cost tool for measuring the effect of driving and secondary task demands on attention. In a Detection Response Task, the participant is asked to drive a vehicle (primary task) and to respond to a frequent and randomly occurring stimuli as quickly as possible (secondary task) (Jahn, Oehme, Krems, & Gelau, 2005). Based on the reaction time to the stimuli, conclusions about the visual and/or mental workload can be drawn as DRT reaction time and miss rates will be elevated with increasing cognitive workloads (Harbluk et al., 2013). The Detection Response Task (DRT) is based on the idea that cognitive interference occurs when several tasks place concurrent demands on cognitive control and there is insufficient amount of resources available to support both tasks simultaneously at the time of demand. The DRT is sensitive to the attentional effects of cognitive interference when another task demands cognitive control (ISO/DIS 17488, 2015).

ISO standard, ISO/DIS 17488 (ISO/DIS 17488, 2015) has been developed to help guide car manufacturers and aftermarket device makers on the use of DRT to assess the cognitive workload of an in-vehicle interface. This standard specifies that the DRT stimuli, which can be either visual (LEDs), tactile (vibration motor), or acoustic (blip), appear randomly every 3–5 s. The participant responds to the stimuli by pressing a micro-switch button. The US Dept of Transportation–NHTSA has recommended the use of a Detection Response Task (over the Lane Change Task) given its sensitivity to changes in the cognitive load of the auditory/vocal memory-scanning task (Ranney, Baldwin, Vasko, & Mazzae, 2009).

The ISO standard recommends the tactile detection response task (TDRT) for examining VCS (ISO/DIS 17488, (ISO/DIS 17488, 2015)) as compared to a visual or acoustic DRT. A light stimuli may exhibit problems with visual eccentricity. It is difficult to determine if a failure to detect the LED light stimuli is due to the driver being cognitively overloaded performing a secondary task or the driver just happens to blink or momentarily look away from the LED light. An acoustic stimuli can be masked by ambient background noises that exist while driving. The tactile DRT (TDRT) resolves visual eccentricity and ambient background noise issues. There is evidence suggesting that the tactile stimuli (TDRT) is more sensitive in detecting differences in cognitive workload compared to light stimuli (Engström, Aberg, Johansson, & Hammarback, 2005). Ranney et al. showed that the TDRT had a higher level of test-retest reliability than visual DRT (Ranney, Baldwin, Smith, Mazzae, & Pierce, 2014).

Numerous studies used the n-back task, which is a working memory numbers recall task, to determine if the DRT is a suitable measure of cognitive workload in the distracted driving domain (Bengler, Kohlmann, & Lange, 2012; Engström et al., 2005; Merat et al., 2006; Ranney et al., 2014; Schindhelm & Schmidt, 2015; Young, Hsieh, & Seaman, 2013). In the n-back task, study participants have to store and recall single digit numbers in working memory that occurred *n* trials ago. Cognitive workload is manipulated by having participant recall a number that occurred two trials ago (2-back) vs zero (0-back) or one trial ago (1-back).

These studies established the DRT as a sensitive measure to capture changes in cognitive workload and have primarily used the n-back task as a proxy for a VCS interaction. Only a few studies have actually tested DRT in conjunction with a VCS (Harbluk et al., 2013; Strayer, Turrill, Coleman, Ortiz, & Cooper, 2014; Strayer, Cooper, Turrill, Coleman, & Hopman, 2015). The n-back task fails to simulate the complexities that a user often encounters while using a VCS while driving. The systems are often imperfect given ambient noise, acoustic similarity of commands, and length of the spoken words. It has been shown that these voice recognition errors can lead to decrements in driving performance (Gellatly & Dingus, 1998). Over 50% of voice control navigation tasks resulted in an error according to a contextual interview study conducted by Wu, Chang, Boyle, and Jenness (2015). Kun, Paek, and Medenica (2007) observed larger steering wheel angle variance with low speech recognition accuracy. McCallum, Campbell, Richman, Brown, and Wiese (2004) observed higher number of collisions in a lead vehicle driving simulator study when a VCS operated at 56% accuracy. Slow system response times along with poor recognition accuracy can push users towards manual input (Ginosar & Hearst, 2014), and recognition accuracy and latency in responding to voice command–may be critical to proper implementation of the technology (McCallum et al., 2004).

There are two primary purposes for this study: (1) to determine if TDRT is a suitable method to capture cognitive workload while driving, and (2) to identify if there are flaws with proxy tasks with respect to capturing workload in systems that are not perfect. Both hypotheses are being examined because there is a breadth of work associated with TDRT and the use of proxy tasks. However, the studies thus far have been conducted on systems that are expected to be perfect. System imperfections however can impact driver performance and the theoretical implications associated with this are examined in this study.

2. Method

2.1. Subjects

The experiment was approved by the University of Washington institutional review board (IRB #45851). There were 48 participants recruited for the study using emails and online classified ads. The participants were evenly distributed across