Driving Assessment Conference   2017 Driving Assessment Conference

Jun 28th, 12:00 AM

A Link Between Trust in Technology and Glance Allocation in On-Road Driving

Claudia Geitner
WMG, University of Warwick, Coventry, UK

Ben D. Sawyer
The Massachusetts Institute of Technology, AgeLab, Cambridge, MA

S Birrell
WMG, University of Warwick, Coventry, UK

P Jennings
WMG, University of Warwick, Coventry, UK

L Skyrypchuk
HMI Research Team, Jaguar Land Rover, Coventry, UK

See next page for additional authors

Follow this and additional works at: https://ir.uiowa.edu/drivingassessment

Presenter Information

Claudia Geitner, Ben D. Sawyer, S Birrell, P Jennings, L Skyrypchuk, Bruce Mehler, and Bryan Reimer

A LINK BETWEEN TRUST IN TECHNOLOGY AND GLANCE ALLOCATION IN ON-ROAD DRIVING

C. Geitner[1], B. D. Sawyer[2], S. Birrell[1], P. Jennings[1], L. Skrypchuk[3], B. Mehler[2], & B. Reimer[2]
[1]WMG, University of Warwick, Coventry, UK
[2]The Massachusetts Institute of Technology, AgeLab, Cambridge, MA, USA
[3]HMI Research Team, Jaguar Land Rover, Coventry, UK
Email: bmehler@mit.edu

Summary: This paper examines whether there is an association between pre-exposure trust in technology and subsequent glance behavior when interacting with a technology that was relatively novel for the majority of participants. After rating their level of trust in technology on a questionnaire, participants drove one of two vehicle models on a highway and engaged in a voice-based navigation address entry task. Subjective ratings of trust in new car technologies were found to be significantly positively correlated with a higher frequency of glances across all coded glance regions during the task. In one of the voice-interface implementations, these higher ratings of trust were also associated with a higher frequency of glances to the user interface, but with fewer long duration (>2s) glances per minute. A lower trust in technology in general showed some association with taking more time to complete interactions. The findings are discussed as highlighting the potential value of further research into the associations between trust and visual scanning behavior.

INTRODUCTION

Trust in technologies can be associated with the patterns of their use (Walker, Stanton & Salmon, 2016). Trust involves the formation of reliance on another actor, person or device, to fulfill a given role for the trustor (Lee & See, 2004), and the level of trust in a technology naturally influences how people interact with it. For example, a low level of trust in a website can lead to reduced attention and engagement (Leichtenstern et al., 2011). Because people respond to technology socially (Nass & Moon, 2000; Lee & See, 2004), behavior changes reflecting the relationship between users and novel technologies are hardly unexpected. Humans also generalize trust, and prejudices may delay or prevent its' formation to social groups (Brewer, 1999). The relative stability of automotive technologies through the latter part of the 20th century, as compared to the explosion of new technologies seen in the past decade, means that vehicle technology can be intuitively segmented into established and novel technologies. The present effort evaluated trust in these "new" categories, as well as technology in general.

When drivers interact with an in-vehicle device, they divide their glances between driving related locations, such as the road and mirrors, and the device (Reimer et al., 2014). During glance time off-road, drivers may miss events on-road. A single eye glance off the forward roadway for longer than 2 seconds while interacting with an in-vehicle device is considered safety critical (NHTSA, 2013). Looking away from the forward roadway while driving contributed up to 80% of "crash events" in a naturalistic driving sample (Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006). Indeed, in-vehicle device design leaves a unique "fingerprint" upon strategic glance behavior (Lee et al., 2017). Therefore, it is not only the fact that drivers looked away from the

road to a device that contributes to driving detriment, but specifically whether the information lost in those glances exceed the demands imposed by the stochastic environment of the roadway. Drivers' allocation of glances is dependent on many factors, such as visual demand of the device and the driver's willingness to engage with the device. In fact, the decision whether a technology is used at all can be influenced by the driver's trust (Walker & Stanton, 2015). Trust is associated with expectations and intentions to use a new technology (Lee et al., 2015). Once a driver has made the decision to engage with a technology, trust has been shown to modulate the level of engagement, e.g. higher engagement in more reliable systems (Walker & Stanton, 2015). What has not been shown is evidence as to whether these relationships influence driver visual behavior during interactions. Understanding these human factors might provide evidence that trust calibration must be considered in the engineering process to avoid or minimize diverting the driver's glances off-road.

The present paper assesses the relationship between self-reported trust relative to three general categories of technology and visual glance behavior during subsequent interaction with a relatively new in-vehicle technology. This work is an exploratory secondary analysis of a large scale study described in Mehler et al. (2016), analyzing a set of metric relationships that has not been reported on previously. The results shed light on the role of self-reported trust in the three categories of technology in the division of visual attention across a range of glance regions as well as specifically to a novel in-vehicle device. It was generally hypothesized, based upon the previous literature, that trust in technology in general, and new vehicle technologies in particular, would influence glance behavior by impacting the frequency or duration of device monitoring.

METHOD

Overview of Data Source, Participants and Analytic Approach
An on-road study was conducted using a 2013 Chevrolet Equinox equipped with the MyLink infotainment system and a 2013 Volvo XC60 equipped with the Sensus system (Mehler et al., 2016). Participants were randomly assigned to either the Volvo or the Chevrolet. This analysis considers 80 participants, equally distributed across the two vehicles over four age groups (18-24; 25-39; 40-54; 55 and older; 20 participants in each group) and by gender (40 females, 40 males). Prior to being exposed to the vehicles, participants completed a questionnaire that included questions concerning trust in technology (detailed below). Participants later engaged in phone contact calling and address entry tasks using the embedded vehicle interfaces and a mounted smartphone. Correlational analyses were employed to examine the extent to which pre-exposure reported trust in technology were associated with later interaction with voice-based address entry into the embedded vehicle navigation systems, specifically task completion time and measures of glance behavior. Data was analyzed in R (Core Team, 2014).

Pre-exposure Trust Questions
Three self-report trust ratings were collected from a broad pre-study questionnaire, each rated on a 10-point rating scale (1 (very distrustful) – 10 (very trustful)):

1. "How would you rate your overall level of trust in technology?"
2. "How would you rate your level of trust in established car technologies (e.g. anti-lock brakes, automatic transmissions, air bags, etc.)?"

3. "How would you rate your level of trust in new technologies that are being introduced into cars?"

Glance Metrics

Measurement of glance patterns is important even in voice interaction tasks. Sawyer et al. (2016) and Reimer et al. (2014) showed that despite reduced glances off-road relative to use of a visual-manual human-machine interface (HMI), voice-based interactions with HMIs often have a significant visual component. Four glance metrics were considered, as detailed in Table 1. Values were calculated for each participant and then averaged. "Device" refers to the center stack region where the visual component of the HMI was located in the vehicles studied.

Table 1. Eye Metrics for the Analysis.

Eye-metric	Calculation
Percent of glance time to the device	Percentage of the total time glancing to the device (center stack region) for the tasks navigation entry 1 and 2. Calculation of the mean from both tasks.
No. of glances per minute	Number of glances (transitions across 9 coded glance regions, e.g. forward roadway, rearview mirror, device, left & right mirror/window, etc.) over the period of a task, for the tasks navigation entry 1 and 2, divided by time. Calculation of the mean from both tasks.
No. of glances to the device per minute	Separate calculation of the number of glances to the device over the period of the task for navigation entry 1 and 2. Calculation of the mean from both tasks.
No. of glances >2s to the device per minute	Separate calculation of the number of glances to the device that were >2s over the period of the task for the tasks navigation entry 1 and 2. Calculation of the mean from both tasks.

The first three glance metrics were selected to reflect characteristics of task switching behavior. The final metric, glances >2s, is common in the literature and utilized in the assessment of driver distraction by the National Highway Traffic Safety Administration (NHTSA, 2013).

Procedure

Before going on-road, participants were trained in a parking lot on the use of the HMI systems, repeating tasks until they felt comfortable to proceed. In an initial driving phase of about 30min, participants familiarized themselves with the vehicle. The voice based navigation address entry task was conducted with three addresses while driving on a divided interstate. The first two addresses were the same for every participant, whereas the third address was the participant's home address which differed between participants. Only the first two addresses are considered in this analysis. The driver's face was video-recorded at 30Hz. This recording was used to manually code the driver's eye glances, as described in Reimer et al. (2014).

The Chevrolet voice interaction system required the address to be entered as a single continuous entry. In contrast, the Volvo system broke up the address into several small steps, comprising: city name, street, and house number. From each step it was possible to return to a previous step. For further detail, see Mehler et al. (2016).

RESULTS

On the pre-exposure questionnaire, participants were asked to rate their previous experience using in-vehicle voice command technology. The majority (65%) reported having never used an

in-vehicle voice interaction technology before taking part in the study; 28% of the participants did not have any experience with voice interaction technology at all. Descriptive statistics for the ratings of the pre-exposure three trust questions (Table 2) showed a similar pattern across the participants assigned to each vehicle, with trust in established car technology nominally higher than trust in new car technology, and general trust in technology falling in-between. On average, the trust ratings across the three questions were markedly skewed to the positive side of the 1 to 10 rating scale. Paired t-tests did not reveal any significant differences in the trust ratings between the two participant groups prior to experiencing the technology in the respective vehicles (all $p > .05$).

Table 2. Distribution of Trust Ratings for the Volvo and Chevrolet Voice Interaction Navigation Systems.

	Volvo		Chevrolet		t-test
	Mean rating	SD	Mean rating	SD	
Trust in technology (general)	7.9	1.4	7.6	1.4	t(77.85)= 1.19, p=0.24
Trust in established car technology	8.8	1.0	8.5	1.3	t(73.25)= 1.14, p=0.26
Trust in new car technology	7.6	1.5	7.5	1.3	t(76.9)= 0.08, p=0.93

The relationship between subjective ratings of trust in technology prior to on-road exposure to two voice-based in-vehicle interfaces and eye metrics during interaction with those interfaces, was analyzed pairwise for each trust item and eye metric with the Pearson product-moment correlation coefficient. Comparisons were conducted separately for the two in-vehicle systems (Table 3). Pre-exposure subjective ratings for trust in technology in general showed only one association of significance (an inverse relationship between total task time and trust in the Chevrolet sample), but no other statistically significant relationships for glance division between road and device. No statistically significant associations were found between ratings of pre-exposure trust in established car technologies and any of the eye metrics.

Table 3. Correlations between Trust Ratings and Observed Behavior (* p <= .05, ** p < .01, *** p < .001).

		Trust in technology (general)	Trust in established car technology	Trust in new car technology
Volvo	Total task time	-.15	-.04	.12
Chevrolet	Total task time	**-.31***	-.1	-.15
Volvo	% of glance time to device	-.08	-.16	.2
Chevrolet	% of glance time to device	-.17	.15	.09
Volvo	Glances/minute	.21	-.09	**.47***
Chevrolet	Glances/minute	.22	.13	**.42***
Volvo	Glances/minute to device	.13	-.07	**.43***
Chevrolet	Glances/minute to device	-.12	.07	.08
Volvo	Glances >2s/minute to device	-.3	-.17	**-.33***
Chevrolet	Glances >2s/minute to device	-.12	-.01	.06

The strongest associations between pre-exposure subjective ratings and eye metrics were found for ratings of trust in new car technology, with the subsequent number of glances per minute showing a strong positive correlation with the degree of pre-exposure self-reported trust in both groups (*p<.01*). This relationship is shown in more detail in Figure 1.

Differences between the groups in the extent to which a relationship appears between pre-exposure trust in new technologies and subsequent glance behavior during interaction with the voice interface appears for the metrics: number of glances per minute to device and number of long duration glances (>2s) per minute to device. In the Volvo interface, levels of trust in new car technologies were found to be significantly correlated with these eye metrics. Higher trust in new car technology prior to the on-road assessment was associated with a higher number glances to the device per minute (*p<.01*), but also with fewer long duration glances (*p<=.05*). In contrast, no significant relationships between trust in new car technology prior to the on-road assessment and to device glance interactions appear for the Chevrolet voice interface.
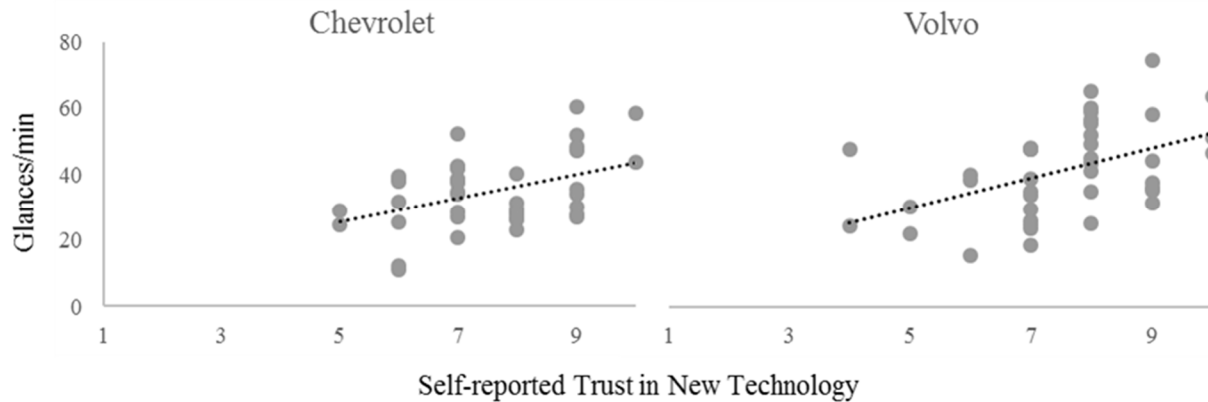

Figure 1. Ratings of trust in new car technology and number of glances per minute aggregated across all glance regions were found to positively correlate (see Table 3) in both vehicles.

DISCUSSION

The results align to some extent with the idea that trust in technology may be associated with subsequent user behavior when engaging with a relatively new form of in-vehicle HMI. During voice-based address entry in both vehicle's embedded navigation systems, previously self-reported higher levels of trust in new car technology were associated with a greater number of glances per minute across the aggregate set of coded glance regions (Figure 1), which included locations such as the mirrors and instrument cluster in addition to glances between the forward roadway and the center stack (where the visual component of the HMI was located). Interestingly, when only glances from another region to the center stack are considered, the number of glances per minute and the pre-exposure trust in new car technology rating were significantly related only in the Volvo sample. In this group, higher trust in new car technology was additionally significantly associated with fewer long duration glances. Together, the observed associations in the Volvo between high trust in new car technologies and eye metrics suggest a higher rate of overall information sampling across glance regions, and a multitasking strategy applying more frequent glances, but of shorter duration (<2s), to the HMI device.

The differences observed in the two vehicles are possibly due to differences in system design. For instance, in comparison to the Volvo, average entry time in the Chevrolet system was shorter, but there were a higher number of both user and system errors (Mehler et al., 2016). The Volvo system broke down the task of entering an address into separate steps for city, street, and number, each with a confirmation which likely both contributed to the lower rate of errors and

facilitated error recovery when errors occurred. In contrast, the full address was entered as a single verbal string in the Chevrolet system. If there was an error on the part of the driver or in the recognition capabilities of the system, the whole address needed to be reentered, a potentially frustrating error handling strategy. Low system reliability is known to be associated with lower trust in a technology (Walker et al. 2015, Lee & See, 2004). A separate analysis (Mehler et al., 2017) found that the number of trials in which an error occurred with the Chevrolet system was negatively associated with participants' post-experimental rating of how likely it was that they would recommend to a friend or family member that they consider buying a car with the voice control technology the participant experienced. The error rates and method of error handling in the Chevrolet system may have led to a reduction in trust during the course of interaction, and, in consequence, may have altered the glance pattern associated with pre-exposure trust over the course of engagement with the system.

It can also be observed that there was a significant negative association between pre-exposure self-reported trust in technology in general and the length of time taken to complete the voice-based address entry with the Chevrolet system. (The same trend appears in the Volvo, but is not statistically significant.) It may be that lower generalized trust in technology is associated with less confidence and greater hesitancy in interaction that can translate into longer completion time. Given that the Chevrolet system did not provide the same level of step-by-step guidance through the address entry process as that provided in the Volvo interface, greater hesitation in users generally less trusting and confident in technology in general makes intuitive sense.

Primarily, however, these findings support the idea that further work might profitably explore the impact of trust upon visual attention, and potentially the utility of eye metrics in revealing trust. They are synergistic with earlier attempts to understand how interaction with very new on-road technologies may be measurably impacted by the trust of the participant population (as in Sawyer, Finomore, Calvo, & Hancock, 2014), and not only in terms of the device itself, but in terms of general attitudes toward technology (Lee et al., 2015). Examination of additional datasets seems warranted to establish if the patterns seen here generalize, and it is important to recognize that correlational analyses do not establish causal relationships. The impact of glance strategy, including the presently identified 'more frequent glance switches', on driver awareness of the on-road situations and overall driving quality also would benefit from further study. Enhanced understanding of how and to what extent drivers' trust in technology affects their visual interaction with in-vehicle devices might aid human factors engineers in contributing to the design of safer systems through knowledgably incorporating the trust variable in further optimizing glance behavior.

ACKNOWLEDGEMENTS

REFERENCES

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, *55*(3), 429-444.

Lee, C., Reimer, B., Mehler, B., & Coughlin, J. (2015). User acceptance of voice interfaces in the automobile. *Proceedings of the 59th Annual Meeting of the Human Factors and Ergonomics Society,* 1641-1645.

Lee, J.B., Sawyer, B.D., Mehler, B., Angell, L., Seppelt, B., Seaman, S., Fridman, L., Reimer, B. (2017, in press). Linking the Detection Response Task and the AttenD Algorithm through the assessment of human machine interface workload. *Transportation Research Record*.

Lee, J., & See, K. (2004). Trust in automation: designing for appropriate reliance. *Human Factors, 46*(1), 50-80.

Leichtenstern, K., Bee, N., André, E., Berkmüller, U., & Wagner, J. (2011). Physiological measurement of trust-related behavior in trust-neutral and trust-critical situations. In *Trust Management V*. Springer.

Mehler, B., Kidd, D., Reimer, B., Reagan, I., Dobres, J., & McCartt, A. (2016). Multi-modal assessment of on-road demand of voice and manual phone calling and voice navigation entry across two embedded vehicle systems. *Ergonomics*, *59*(3), 344-67.

Mehler, B., Reimer, B., Lee, C., Kidd, D., & Reagan, I. (2017, June). Considering self-report in the interpretation of objective performance data in the comparison of HMI systems. *Proceedings of the 9th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design.*

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81-103.

NHTSA (2013). *Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices*. National Highway Traffic Safety Administration Pub. No. NHTSA-2010-0053.

Reimer, B., Mehler, B., Dobres, J., McAnulty, H., Mehler, A., Munger, D., & Rumpold, A. (2014). Effects of an 'Expert Mode' voice command system on task performance, glance behavior & driver physiology. *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications,* ACM.

Sawyer, B.D., Mehler, B. & Reimer, B. (2017, June). An antiphony framework for dividing tasks into subtasks. *Proceedings of the Ninth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*.

Sawyer, B., Finomore, V. Calvo, A., & Hancock, P. (2014). Google glass a driver distraction cause or cure? *Human Factors, 56,* 1307–1321.

Sawyer, B. D., Lee, J., Dobres, J., Mehler, B., Coughlin, J. F., & Reimer, B. (2016). Effects of a voice interface on mirror check decrements in older and younger multitasking drivers. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*(1), 95-100. Washington DC.

Walker, G. H., Stanton, N. A., & Salmon, P. M. (2016). Trust in vehicle technology. *International Journal of Vehicle Design, 70*(2), 157-182.

Wierwille, W. (1993). Demands on driver resources associated with introducing advanced technology into the vehicle. *Transportation Research Part C, 1,* 133–142.