Check for updates

Autonomous Driving Systems: A Preliminary Naturalistic Study of the Tesla Model S

Mica R. Endsley, SA Technologies, Inc.

Autonomous and semiautonomous vehicles are currently being developed by over 14 companies. These vehicles may improve driving safety and convenience, or they may create new challenges for drivers, particularly with regard to situation awareness (SA) and autonomy interaction. I conducted a naturalistic driving study on the autonomy features in the Tesla Model S, recording my experiences over a 6-month period, including assessments of SA and problems with the autonomy. This preliminary analysis provides insights into the challenges that drivers may face in dealing with new autonomous automobiles in realistic driving conditions, and it extends previous research on human-autonomy interaction to the driving domain. Issues were found with driver training, mental model development, mode confusion, unexpected mode interactions, SA, and susceptibility to distraction. New insights into challenges with semiautonomous driving systems include increased variability in SA, the replacement of continuous control with serial discrete control, and the need for more complex decisions. Issues that deserve consideration in future research and a set of guidelines for driver interfaces of autonomous systems are presented and used to create recommendations for improving driver SA when interacting with autonomous vehicles.

Keywords: driving, autonomy, automation, situation awareness, mode awareness, distraction, interface design

The automotive industry is rapidly developing semiautonomous and fully autonomous vehicles for the driving public. This includes a variety of advanced aids for navigation, collision and blind spot warning systems, and more advanced features such as automated parking, adaptive cruise control (ACC), and automated lane-following systems. Whereas Google is focused on developing a fully autonomous automobile that will require no intervention from the driver, other companies have been rolling out autonomous features more gradually that are intended to augment and aid the driver.

These efforts are not without risk, however. Developing autonomous systems that integrate well with the human driver is considerably challenged by the tendency for automation to put drivers out of the loop (Wickens & Kessel, 1979; Young, 1969), due to reduced situation awareness (SA; Endsley & Kiris, 1995). When monitoring automation, people are often slow to detect that a problem requiring intervention exists, and they can be slow to arrive at a sufficient understanding of the problem to do so effectively.

For example, a Tesla Model S was recently involved in a fatal crash that killed the driver when its autopilot system failed to detect an 18-wheeler that turned in front of the car (Soloman, 2016). Would the driver have been able to detect the truck and stop in time had he not been reliant on the automation? Or would the driver also have failed to see the truck against the bright sky as Tesla claims, making the automation irrelevant? In another accident, a Tesla driver failed to brake, rear-ending another car, because she trusted that the autopilot would do so and she intervened too late to prevent the collision (Gitlin, 2016).

SA is a fundamental precursor to successful performance in many highly dynamic and complex domains, including driving (Endsley, 1995; Gugerty, 1997; Horswill & McKenna, 2004; Ma & Kaber, 2005). Drivers need SA of many factors: the speed, fuel level, and functioning of one's vehicle; the relative distance, speed, and trajectories of other vehicles and pedestrians; the impact of weather and hazards on vehicle safety; one's location on the desired route, distance to the next

Address correspondence to Mica R. Endsley, SA Technologies, Inc., 1910 S. Stapley Drive, #221, Mesa, AZ 85204, USA, mica@satechnologies.com.

Journal of Cognitive Engineering and Decision Making
2017, Volume 11, Number 3, September 2017, pp. 225–238
DOI: 10.1177/1555343417695197
Copyright © 2017, Human Factors and Ergonomics Society.

turn, and projected time and distance to the destination; compliance with posted speeds and other applicable laws; and knowledge of abnormal vehicle equipment states and their effect on vehicle safety and performance (Bolstad, Cuevas, Wang-Costello, Endsley, & Angell, 2008). SA includes a consideration of not only driver attention and perception but also the ability to (1) make meaningful assessments in terms of the significance of perceived information to one's goals (e.g., the need to change lanes in advance of an upcoming exit or accident); and (2) project likely future events to make proactive decisions (e.g., slowing down in response to brake lights down the road or rerouting around a traffic jam).

Poor SA has often been implicated in vehicle crashes (Gugerty, 1997). Two of the leading causes of driving accidents—improper lookout and inattention—are examples of failures to maintain SA (Treat et al., 1979), as well as distraction and recognition errors (e.g., looked but did not see or interpreted incorrectly; Sabey & Staughton, 1975). Ma and Kaber (2005) found that that all levels of SA (perception, comprehension, and projection) appear to have an impact on operational driving behaviors, and Ward (2000) and Matthews, Bryant, Webb, and Harbluk (2001) showed how different types of driving tasks (operational, tactical, and strategic) require all three levels of SA.

THE CHALLENGE FOR OPERATORS OF AUTONOMOUS SYSTEMS

Considerable research has been directed at the challenge of human-automation integration in a variety of environments, including aviation, driving, and process control. Summarizing some 30 years of research on the topic, the human autonomous-system oversight model shows that a driver's ability to intervene and take over from automation when needed depends on his or her SA of critical information in the driving environment (Endsley, 2017). Three main factors will have a large effect on driver SA with autonomous systems.

Attention and Trust

SA is directly influenced by one's level of attention to relevant driving information, which is affected by the level of trust in the automation and the presence of competing secondary tasks and which is mediated by the effectiveness of the vehicle displays. Trust is significantly affected by the capabilities of the automation, primarily its reliability and robustness (Hancock et al., 2011; Lee & See, 2004). When the automation is more reliable and more robust, trust is increased, and people are more willing to divert their attention to competing tasks, such as day dreaming, operating in-vehicle technologies, talking or texting on cell phones, eating, grooming, or performing other extraneous tasks (Carsten, Lai, Barnard, Jamson, & Merat, 2012; Hergeth, Lorenz, Vilimek, & Krems, 2016; Kaber & Endsley, 2004; Ma & Kaber, 2005; Sethumadhavan, 2009), thus directly lowering their SA.

Engagement and Workload

The driver's level of workload and engagement significantly affects SA. Even when drivers are vigilant, they can become much less engaged under automation (more like a passenger than a driver), which lowers SA and slows response to hazardous situations (Endsley & Kiris, 1995; Manzey, Reichenbach, & Onnasch, 2012; Metzger & Parasuraman, 2001). How engaged operators are and how much workload they experience is a function of the way that the automation is implemented—particularly (a) the level of automation (i.e., which aspects of a task are automated); (b) adaptive automation, which intersperses periods of manual and automated control; and (c) the amount of control granularity, determining how detailed guidance needs to be. An automation conundrum exists in which the more automation is added to a system and the more reliable and robust that automation is, the less likely that human operators overseeing the automation will be aware of critical information and able to take over manual control when needed (Endsley, 2017).

Mental Model

Finally, SA is significantly affected by the accuracy and completeness of the driver's mental model of the system. Developed largely through training and experience, as well as through the transparency of the system interface, mental models create a set of expectations that drivers use to interpret the events of the driving environment and the actions of the system. Unfortunately, as automation becomes more capable and more complex, creating an accurate mental model becomes quite difficult. The use of new deep-learning techniques for the development of autonomous systems, creating a frequently changing and oblique set of system logic, will make maintaining an accurate mental model even more challenging (U.S. Air Force, 2015).

A majority of research on automation in real-world settings comes from experiences and accidents in the aviation and process control industries over the past 30 years. This research forms an important basis for understanding the challenges that drivers will face in autonomous vehicles in the future. In comparison, the driving environment may see even more significant problems with autonomous features in that cars operate much closer together than aircraft, leaving very little room for error in many traffic scenarios, and drivers lack the high level of selection and training that is common among commercial pilots, who operate the most sophisticated automated flight management systems.

RESEARCH ON AUTOMOBILE AUTONOMY

Over the past few years, a number of studies in simulation environments have focused on determining how new autonomous systems will affect drivers. Petermeijer, Abbink, and de Winter (2015) investigated automated steering systems and found that they improve driver satisfaction and performance but also increase the time to recover from a system shutdown, demonstrating an out-of-the-loop problem.

Research also indicates that driver distraction is likely to increase with the implementation of autonomous systems. Visual monitoring frequency and duration were found to decrease as trust increased during a highly automated driving task (Hergeth et al., 2016). De Winter et al. (2014) found a 261% increase in drivers performing other secondary tasks with highly automated driving. Carsten et al. (2012) showed that drivers in highly automated vehicles (with both ACC and automated lane following) increasingly engaged in secondary tasks, leading to significant workload spikes when interventions were needed, often exceeding driver capacity. Similarly, Merat Jamson, Lai, and Carsten (2012) discovered that when drivers using autonomous systems redirected their attention to secondary tasks, they were much slower in responding to critical incidents, and this effect was much worse in periods of underload. In addition, interacting with the automation itself creates a distraction adding to the problem (Neubauer, Matthews, Langheim, & Saxby, 2012).

Not all research has shown that automation decreases SA, however. Beller, Heesen, and Vollrath (2013) found improved SA and collision response time when the reliability of the automation was displayed to the driver. Ma and Kaber (2005) also found improved SA with ACC, which they credited to an increase in drivers' available mental capacity for attending to traffic. However, this benefit disappeared when the drivers used cell phones, confirming the negative effects of dual tasking while overseeing automation. Interestingly, it was the driver's ability to project future events (Level 3 SA) that was most affected. As talking on cell phones has been shown to tax the central executive function of working memory and impair driver SA (Hennan, Herdman, Brown, & Robert, 2014), it is likely that the drivers had insufficient capacity for projection, an important aspect of driving, as well as lowered engagement.

These studies confirm many concerns about the effects of autonomous driving systems. However, most studies were conducted in driving simulators on tasks of a relatively short duration, which may not reflect real-world conditions. Strayer and Cooper (2015), for example, showed that the distraction level of many tasks was much higher in real-world settings as compared with simulator studies. In addition, drivers' behavior may change over time as they become accustomed to the automation, learning when to trust it and adapting their driving strategies, or as they become increasingly bored with a less demanding and engaging task.

Although many automobile companies are actively engaged in testing new autonomous features, many of these have not yet been widely fielded. Tesla has been taking a more aggressive approach, rolling out advanced automation features in its commercially available vehicles.

Figure 1. Main automation driving display.

Therefore, I elected to conduct a naturalistic driving study in a Tesla Model S, carefully documenting my experiences over a 6-month period.

Although such an approach lacks the experimental control that one finds in simulator studies, this naturalistic driving study provides a unique view on the challenges that drivers may face in dealing with new autonomous automobiles in realistic driving conditions, without the limitations of simulation studies. It should be noted that this analysis reveals only my own experiences, and I analyzed the system through the lens of what is known about human-autonomy interaction and system design (as compared with typical research studies employing multiple naïve subjects). This analysis forms an early look at how autonomous driving software may affect driver SA and performance, to determine how well previous research on human-autonomy interaction applies in the driving domain, the types of issues that deserve consideration in future research, and recommendations for improvements in the design of system interfaces for automobile autonomy.

AUTOMATION SOFTWARE TESTED

The study was conducted in a Tesla Model S 70. The car's autonomy software was at version 7.0 at the beginning of the study. It automatically received a total of five software updates (7.1–7.1.2.18) over the 6 months of the study. The following automation software was tested, and Figure 1 shows the displays provided to the driver.

Adaptive cruise control: The ACC kept the vehicle at a speed set by the driver. If it approached the rear of another car going slower, the ACC would slow to keep the car at a set distance behind the car in front (5 car lengths).

Autosteer: A part of Tesla's autopilot, the autosteer feature was in beta mode during this study. Autosteer keeps the car on a track along the center of the lane markings in the road or a set distance from the side if it senses only one lane marking. Autosteer was always used with ACC, although ACC could be used alone.

Auto lane change: While in autosteer, the auto lane change (ALC) automatically moves the car into the adjacent lane when the driver turns on the turn signal. The driver is responsible for making sure that the lane is clear before activating the ALC.

Navigation system: A GPS-based map navigation system provided automatic navigation to any destination entered. The driver was responsible for following the navigation directions, as the autosteer function was not directly tied to the map.

Summon: The summon mode was introduced in the first monthly update, approximately 6 weeks into the test period. Summon allows the driver to move the car into and out of a garage or parking place while outside the vehicle, using the key fob or smartphone. The driver is responsible for making sure that the car is clear of obstacles and for starting and stopping the vehicle.

Warnings: The current speed limit was displayed on the main driving display (as updated by sensors when a legal speed limit sign was passed), along with a temporary increase in size when the speed limit was exceeded by >5 mph. In addition, a side collision warning system displays fans alongside the vehicle to show areas where it detects an obstacle or another vehicle. Lane departure warnings were provided via haptic feedback when the vehicle crossed a lane marking without the turn signal on. Emergency collision braking was added during a software update, providing automatic braking if an imminent collision is sensed.

STUDY METHOD

After picking up the Tesla Model S and receiving the normal instruction on its functioning from the service representatives, I drove 200 miles as an initial training period. My goal during the 6-month period of the study was to use the automation whenever it was safe to do so during all normal driving. I collected data on my experiences with the automation via an end-of-trip questionnaire that included the length of the trip, the driving conditions, the automation used, any problems that I encountered during the trip, and a rating (on a 5-point scale) of the satisfaction level with, trust in, and usefulness of the automation, along with a subjective workload rating and other comments.

I also collected data on my SA while driving. Bolstad et al. (2008) discussed the applicability of SA measures to the driving task. They recommended direct objective measures, such as the Situation Awareness Global Assessment Technique or real-time SA probes, which provide detailed information on all three levels of SA, as the most sensitive and diagnostic. These measures are typically collected in simulator settings, which allow for easy data collection. In this real-world driving task, I used a modified technique in which at random times I verbally recorded my response to a number of SA probes during the driving task (rather than during freezes in the simulation, which the usual case with the Situation Awareness Global Assessment Technique).

A timer was set to chime at a random point between 3 and 10 minutes. At that time, I quickly memorized my answer to each of five SA probes, without looking around or at my displays. When it was safe to do so, I recorded my answers to the probes and the correct answer on a voice recording that was later entered into a computer. The SA probes were selected from a set of SA requirements across multiple goals associated with driving (Bolstad et al., 2008):

- What is your trajectory as compared with the car ahead? *Gaining on, receding from, maintaining following distance, not applicable*
- Where are surrounding vehicles? *Left side, left rear, rear, right rear, right side, none*
- Are you within a half mile of the next turn? *Yes, no*
- What speed are you going? *Within 5-mph blocks*
- Are you within 5 mph of the speed limit? *Yes, no*

(I did not collect SA data on any trip in which a passenger was present in the automobile.) In addition, at that time, I recorded my workload via the subjective workload assessment technique via a 1- to 3-point rating on the subscales of time, effort and stress, my level of engagement (9-point scale), and the automation currently in use.

For purposes of comparison, I also collected the SA data in a control condition for 1 month prior to starting data collection on the Tesla. In the control condition, I drove a Lexus 350ES that had no advanced autonomy features. It was equipped with a traditional cruise control and an in-vehicle GPS-based map navigation system, which I did not use during the data collection period.

FINDINGS

I completed a total of 2,684 miles in the 6-month period of the test, >106 trips with an average trip length of 25.6 miles: 78% of the trips involved highway driving, 45% city driving, and 8% country driving (many trips were based on combined types). Some form of automation was used in 84% of the trips during the study period: navigation guidance (50%), ACC

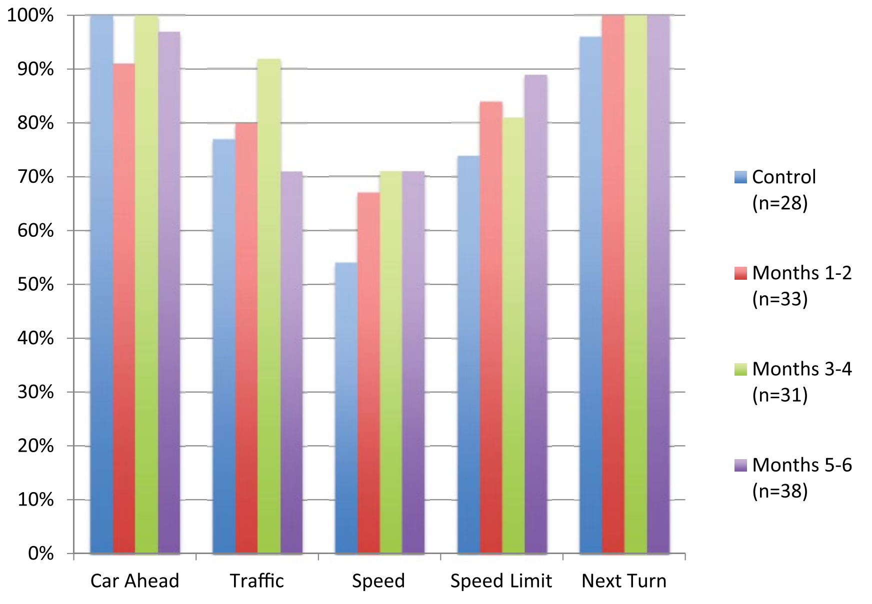

Figure 2. Situation awareness across the period of study.

(82%), autosteer (82%), ALC (72%), and summon (4%).

An analysis of my SA, as measured during driving, did not show significant changes ($p < .05$) over the period of the study (Figure 2), potentially due to the relatively small sample size, with only one driver. Trends show that awareness of my relation to the car ahead and being within a half mile of the next turn remained high throughout the study, most likely because there was little variation on the answers to these questions. Knowledge of my speed and speed limit conformance increased with the automation, primarily due to the consistency of ACC as opposed to manual control of speed; however, this increase was not statistically significant.

Awareness of other traffic increased somewhat initially but then decreased in the last period of the study. In general, when the autosteer and ACC were engaged, I found it quite easy to look around more and observe adjacent traffic. Over time, however, as boredom and vigilance challenges set in, my attention wandered to competing tasks, such as daydreaming, adjusting the navigation system and sound system, or attending to text messages.

I should note that collecting SA data in this study required that I resist looking at the relevant information while determining my responses (generally 5–6 seconds) and only afterward looking around to see if my answers were correct (typically another 1–2 seconds); I would then rapidly record those answers and a correct/incorrect score from memory via a voice recorder while driving, usually within 30 seconds. This is more challenging than assessing SA with the Situation Awareness Global Assessment Technique in a simulator, which allows the scene to be stopped and blanked while answering SA probes. My subjective impression was that it most likely interfered with driving while this was occurring. For safety reasons, the use of SA probes in simulators is preferable to attempts to collect SA probe data during real-world driving.

Overall, I experienced an automation-related problem on 30% of the trips. Figure 3 shows the frequency and type of problems encountered with the automation used during the study. Despite the incidence of automation-related issues, my overall satisfaction with, perceived usefulness of, and trust in the automation all increased over the period of study, as shown in Figure 4. These increases were partially due to improvements in the automation that occurred in the monthly software updates, but they were primarily due to my adapting to what the automation could or could

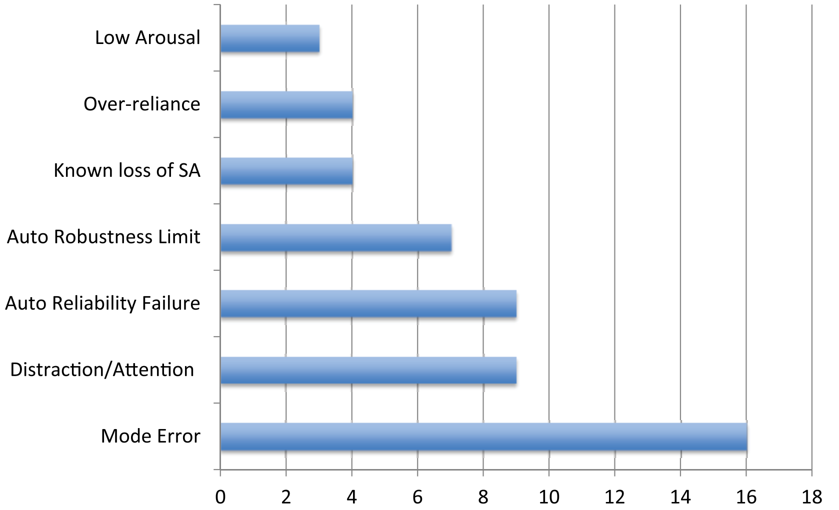

Figure 3. Incidence of automation-related problems. SA = situation awareness.

Figure 4. Mean satisfaction, usefulness, and trust in automation and workload ratings over the period of the study (scale: 0 = *low*, 5 = *high*).

not do and adjusting my usage of it accordingly. My workload was low across the study.

AUTOMATION-RELATED PROBLEMS

Training and Mental Models

Developing and maintaining an accurate mental model of how the automation worked turned out to be quite a challenge. Tesla's service representative provided an initial verbal description of all the controls and displays when I picked up the vehicle, as well as a verbal description of the autonomy features. Although the service representative initially informed me that the automation was 100% reliable, after repeated questioning, a more experienced service representative was brought in (i.e., 3 months on the job, as opposed to 3 days), who not only explained that it was not 100% reliable but also detailed some of its limitations and offered me a test drive. The training experience was rather ad hoc and guided mostly by the questions and knowledge of the driver.

Developing a mental model of how the automation worked was, for the most part, up to the driver. Although some information on how each feature worked is included in the manual, it was not comprehensive and did not explain the details of how each automation feature works, either alone or in conjunction with other features. I therefore had to rely on trial and error to determine how the automation would work in different driving conditions.

In many cases, I was not exactly certain why the automation behaved the way that it did, and I was forced to generate my own assumptions. The challenge of creating an accurate mental model of the system was significantly compounded by the rapid rate of change of the software. Although the driver is notified of each update and an update release note is provided on the in-vehicle display, these were often insufficiently detailed to explain many of the changes. Furthermore, the update notes were not available online, so I was restricted to reading them only when in the vehicle.

For example, the autosteer would not initially work on sharp curves. About 4 months into the test, I discovered that it had learned to handle these curves. The software had learned to slow down going into a curve and was improved at tracking the lane markings. This was surprising, however, as none of the update release sheets mentioned this change.

The summon mode was added to the system in a software update, with only written instructions available. In a subsequent update, its operation changed significantly (via a smartphone app rather than the key fob) with a very confusing explanation provided. It took numerous attempts through trial and error to figure out how to use this new feature.

Mode Awareness and Mode Confusion

Mode confusion was the most frequent problem I encountered. In the majority of cases (n = 11), this confusion stemmed from the fact that the lever controlling the ACC and autosteer functions was located directly below the turn-signal lever on the left-hand side of the steering column. The turn signal provided traditional signaling and activated the ALC when the autosteer was on. Pulling the cruise control stalk forward one time activated the ACC and two times, the autosteer. A frequent error was to accidentally perform this action on the turn-signal lever above it, which would instead flash the brights. Moving the cruise control stalk up or down increased or decreased the speed setting. Frequently, however, I accidentally performed the desired action on the lever above, turning on the turn signal instead, or vice versa (i.e., changing the speed setting when I meant to activate the turn signal). This error led the car to speed up on several occasions when I intended to change lanes, necessitating an intervention.

In two cases, I thought that the automation was on when it was not (the autosteer failed to capture on activation for unknown reasons), and in three cases, I did not realize that the ACC was still on after I took over manual control. When the driver presses the brakes, the ACC and autosteer will turn off; however, if the driver turns the steering wheel, only autosteer is canceled, and the ACC remains on. When I turned the wheel to exit the freeway at an off-ramp, for example, I was surprised that I was still traveling very fast and needed to brake as well to disconnect the ACC. This problem also occurred on a sharp curve when I took over manual steering; I was surprised that the ACC was still engaged and that the car was going too fast for the curve.

Unanticipated Mode Interactions and Emergent Behaviors

Initially, I was instructed that the various automated systems (ACC, autosteer, and navigation) all worked independently. However, I found that the logic of the system created some unusual and unexpected emergent behaviors, which was further compounded by software updates that linked the behavior of the modes in certain circumstances. For example, when I was behind a car that was going slower, the ACC would slow down to maintain a set following distance. When I changed lanes, however, the car would unexpectedly surge to regain the set ACC speed.

As a safety precaution, when both the ACC and the autosteer are active, the vehicle's speed is limited to no more than 5 mph over the speed limit on roads without a divider. If the driver takes over manual steering control or if the road later becomes divided (e.g., a median is present), the car can suddenly surge to the higher ACC speed setting, creating some unexpected experiences.

Automation Reliability and Robustness

The ACC generally performed reliably and maintained the set following distance behind cars in front of it. As the car in front slowed, so would my car, generally negating the need to brake when coming to stop lights. If I was approaching fully stopped traffic or there was no car in front of me, however, I did need to brake for stoplights. This created a more complex dual-decision situation. Instead of always braking at a red light, I needed to make two decisions—"Is a stop needed?" and "Do I need to intervene?"—which can increase response time.

The autosteer function was very reliable on straight highways, hewing to a steady track. It was not sufficiently robust in different situations, however. It could not be easily used on surface streets, as it did not know what to do in intersections where line markings were not present. As the automation generally quit at this point, I stopped trying to use it on surface streets for most of the test period. It also could not handle sharp turns, at least initially, although its performance improved over the course of the test period.

On several occasions, the automation quit unexpectedly, returning to manual control when in the middle of a curve, which was quite frightening. It also had trouble in handling lane merges; it did not know what to do when the lane split into two lanes. More important, it did not react to a reduction in lanes that led to another car merging into it.

The release notes state that autosteer does not work well when lane markings are absent, faded, or ambiguous or in rainy, snowy or foggy conditions. In my experience, it lost the markings only once on a country road with very poor line markings. During a hard rainstorm, it lost track of the road markings only when they were no longer visible to me as well.

The ALC function worked very reliably. It changed only one lane when activated each time, requiring a second activation to change lanes again. Because the Tesla's sensors could not see very far ahead or behind the vehicle, manual check of the adjacent lane was always required to ensure that there was no approaching traffic. Awareness of vehicles directly next to the Tesla or in its blind spot appeared to be inconsistent. On at least one occasion it prevented me from changing lanes into an adjacent vehicle, and on another occasion permitted the lane change which then required manual intervention.

Use of ALC in moderate or heavy traffic was quite difficult. It required many discreet changes of speed to align the car with an open slot, consistent with the speed of adjacent traffic, before changing lanes. I found it much easier to perform lane change and merging maneuvers manually.

The summon mode worked well, moving slowly in and out of the garage and stopping when commanded. When an obstacle was placed in its path, the car tried to maneuver around it, leading it into an odd angle in a small space. I used this feature only a few times during the test, to move out of a parking space when someone parked too close, as there was little other need for it.

SA, Distraction, Arousal, and Overreliance

In some ways, I found that my SA increased when using automation. Instead of being highly focused on maintaining vehicle speed and trajectory, I was able to look around more at other traffic, signage, and information displays.

I also found several instances of distraction and loss of attention that were concerning, however. Over time, my attention wandered, and I found that I was much more likely to daydream, interact with the navigation system or sound system, and even text while driving, which I never did when driving manually. The highly reliable automation was quite seductive, serving as an enabler of bad behavior.

As I generally conducted such extraneous tasks only in situations where the automation was highly reliable, this might not be a problem. However, when the unexpected occurs (e.g., a lane merge or unexpected curve), that temporary loss of SA could be lethal. Thus, even if SA is the same on average as under manual control versus automation, it can become more variable and prone to highs and lows, with a consequent risk for accidents.

An even more insidious problem was a loss of engagement while using the automation. When driving the Tesla in autopilot mode, I found that I was surprisingly slow to react when my car and a recreational vehicle next to me began to come closer together as our lanes on the highway merged. It took extra seconds to realize that the automation was not going to handle the situation. I was able to take control in time, but this might not always be the case

In one situation, I became quite drowsy while driving late in the evening on autopilot. On the positive side, the autosteer kept me safe by keeping the car in the correct lane and eliminating the risk of going off the road. However, had it also enabled me to fall asleep, I would have been unaware of the need to intervene if required.

One of the biggest challenges I found was in how Tesla handled unanticipated automation transitions—when the automation could not handle a situation and unexpectedly returned the function to manual control. The transition was signaled by a low audio tone and a change in color of the automation mode symbol on the display from blue to white. This happened at least eight times during the study. Unfortunately, the audio tone provided is not nearly loud enough or salient enough to catch the driver's attention, particularly over the sound system, and may not be perceived by a driver who is already distracted or sleepy. Furthermore, no advanced warning was provided in any of these cases, largely because circumstances were beyond what the autonomy software or sensors could handle, meaning that immediate action by the driver was necessary.

The Tesla autosteer function requires that the driver keep his or her hands on the wheel at all times, and it informs the driver of this with a warning message whenever it is activated. If the car does not detect hands on the wheel, it will provide a warning signal after 3 to 5 minutes and again if hands are still not detected. If it continues to not detect the driver's hands on the wheel, it will start slowing down and gradually stop. I found that the sensor was poor at detecting my hands on the wheel and that the warning occurred frequently in error, causing significant frustration. (Note: After owning the vehicle for almost a year, I discovered that the wheel does not have a pressure sensor but rather reacts to left-right steering inputs, thus explaining this problem. This is another example of undocumented functionality leading to a poor mental model.)

The biggest limitation is that, although this feature is likely to provide a good backup for the case where a driver has fallen asleep, it is inadequate for handling the loss of SA that can occur due to low arousal, distraction, or loss of engagement. Having one's hands on the wheel is not the same as having one's mind on the road.

RECOMMENDATIONS

Although this was a preliminary study based on my experiences over a 6-month period, I found many significant issues that deserve further investigation with a larger subject pool and with additional autonomous vehicle designs to determine their generalizability. Such studies could employ simulators where more stringent data collection is possible, as well as naturalistic and longitudinal studies that can examine how typical drivers experience interactions with autonomous vehicles under everyday conditions. From just this limited study, however, several recommendations for driver training and vehicle interface design can be made.

Training

The need for more detailed driver training on the functioning of automated systems is critical. Customer service representatives may be very inexperienced or lack the skills for communicating the complexities of automated systems. Drivers will need much more detailed and interactive training on not just how to operate an automated system but also on how it will perform in different driving conditions and how different automated features may interact in practice such that they can form accurate mental models needed to appropriately calibrate their trust and expectations. This training can be easily delivered via online video software, which can be updated as the systems mature.

Driver Displays

Detailed guidelines for improving operator SA have been developed (Endsley & Jones, 2012) that can be applied to driving. Table 1

TABLE 1: Evaluation of the Tesla Interface per Guidelines for Supporting Human-Autonomy Interaction

	Support human understanding of autonomous systems		
1	Automate only if necessary—avoid out-of-the-loop problems if possible		Combined ACC/autosteer creates an out-of-the-loop issue and added complexity in predicting system behavior.
2	Use automated assistance for carrying out routine tasks rather than higher-level cognitive functions		Key decisions regarding route selection and route following are left to driver. Automation of many routine tasks is useful, such as headlight and brights on/off, garage door operation, backing camera activation, and locking/unlocking. Autobraking is inconsistent.
3	Provide SA support rather than decisions		Most decisions are left to driver. Speed limit and lane departure warnings are beneficial. Exceedance of speed limit needs continued display with more salient representation (e.g., red outline). Side collision warnings are not salient.
4	Keep the operator in control and in the loop		Driver is in control of each automated feature. Driver engagement during autosteer/ACC is low. New strategies are needed to incorporate the driver and improve engagement.
5	Avoid the proliferation of automated modes		Multiple modes and mode interactions create complexity and unexpected behaviors. Better integration of mode operations and deconfliction of mode activation/deactivation methods are needed to improve mode operation and awareness.
6	Make modes and system states salient		Good display of current modes. Audio cues of unanticipated transitions to manual control lack needed salience. Unique audio cues are also needed to alert driver to partial mode changes (e.g., autosteer off but ACC still on).
7	Enforce automation consistency		Consistency in the terminology and information placement among modes was good. Some unexpected behaviors arose from mode interactions.
8	Avoid advanced queuing of tasks		Autosteer, ACC, and navigation all set up tasks to be carried out in advance, which create the lowest levels of SA. Approaches that maintain operator involvement in the execution of tasks should be considered.
9	Avoid the use of information cueing		No information cueing provided.
10	Decision support should create human-system symbiosis	N/A	No decision support systems provided
11	Provide automation transparency		Displays of the road and vehicles sensed by the system good for supporting shared SA with driver and provide projection of system actions. Improvements in supporting understandability of system actions and predictability of braking and speed changes are needed.

(continued)

TABLE 1: (continued)

	Minimize complexity of autonomous systems		
12	Ensure logical consistency across features and modes	○	Some unexpected behaviors arose from mode interactions, particularly as situations changed. Modes behaviors should be reviewed and modified to consider such interactions in context to conform to driver expectations.
13	Minimize logic branches	○	System complexity increased as new rules were added to the system. Complexity can be minimized by reducing the linkages and conditional operations contained in the autonomy, thereby avoiding modes with their multiple-branch logic as much as possible.
14	Map system functions to the goals and mental models of users	○	A clear mapping between user goals and system functions should be present—for instance, merging with traffic or exiting highways could be single-step actions, rather than requiring multiple interactions with different modes and an understanding of how those interact.
15	Minimize task complexity	●	Actions to interact with the automation were simple and intuitive.
	Support situation awareness		
16	Integrate information to support comprehension of information (Level 2 SA)	○	Display of lane and vehicles in front is good. Improved display of objects on the sides, blind spot, and rear of vehicle is needed. Improved display of exceedance of speed limits, obstacles below bumper, and stationary vs moving objects on sides is needed.
17	Provide assistance for SA projections (Level 3 SA)	○	Power and range projections are good. Improved display to predict actions of autonomy is needed, particularly those that it cannot handle. Improved display of future road hazards, traffic, and side collision warnings is needed.
18	Use information filtering carefully	◒	Automatic replacement of information displays (e.g., presenting phone call information) covered only low-priority data.
19	Support assessments of confidence in composite data	■	No confidence information associated with system was provided.
20	Support system reliability assessments	■	Trust and effective judgments on when to intervene depend on an accurate assessment of system reliability for performing the task at hand. Interfaces should make explicit how well the autonomy is currently performing and its ability to handle upcoming or contemplated tasks.

■ Very Poor ◪ Poor ○ Average ◒ Good ● Very Good

Note. ACC = adaptive cruise control; SA = situation awareness.

shows how displays of the Tesla Model S compare against these guidelines, as well as areas where they can be improved to help address some of the interaction problems found in this study. In many ways, Tesla provided an excellent visual interface for the autonomy software that was, for the most part, intuitive and easy to use. Displays were clearly presented and easy to navigate.

However, the challenges are significant in terms of creating effective interfaces and methods for supporting drivers' understanding of automation, given its inherent complexity, and for compensating for the very real problems associated with loss of engagement and distraction. Because Tesla is ahead of many other companies in delivering a suite of more highly automated functions to the driving public, it is worth examining how well its approach addresses these challenges. The recommendations in Table 1 provide an outline of issues that all automotive companies will need to address as they launch autonomous features into the marketplace. Furthermore, in that many people drive multiple vehicles, within their families or from rental car companies, there is an impetus to develop consistent standards for driver-autonomy interfaces. These guidelines point the way to many issues that will need to be addressed in these standards. In particular, the need for a high level of automation transparency, providing both understandability and predictability of automation actions, will be critical to address any inconsistencies in automation behavior among vehicles and changes over time.

CONCLUSIONS

This study is limited in that it captures the experiences of only one driver and that the SA probes and assessments were self-administered. However, it serves to confirm many of the research findings from simulation studies on the risks to SA associated with highly automated driving, and it augments research from other in-vehicle studies on automation (Blanco et al., 2015; Blanco et al., 2016). It also provides many insights into the types of challenges with operating highly automated vehicles that drivers can expect to encounter, including mode confusion, mental model development, and support for unexpected automation transitions. In particular, I found that driver training on new autonomy features needed much more attention. Future research studies should further explore these issues with larger subject pools and both simulator-based and naturalistic research methods.

Automation capabilities are improving rapidly, and many of the reliability and robustness problems reported here are likely to decrease or even disappear over time. However, the driver interaction issues will remain and must be addressed. Although this study focused on the Tesla automation software, I should note that most of these issues are quite common across other automotive companies and are not unique to that vehicle.

The coming wave of automation is creating a significant change in the driving experience. As drivers become more passive in the operation of their vehicles, significant increases in nondriving behaviors will become prominent. Much greater attention to the design of the driver interface will be needed to help compensate for this natural human tendency. Simple admonishments and warnings to pay attention will not be sufficient, as demonstrated through years of research on vigilance and monitoring. The success of automobile automation efforts and the safety of the driving public depend on careful attention to system designs that incorporate driver needs.

REFERENCES

Beller, J., Heesen, M., & Vollrath, M. (2013). Improving the driver-automation interaction: An approach using automation uncertainty. *Human Factors*, *55*(6), 1130–1141.

Blanco, M., Atwood, J., Russell, S., Trimble, T., McClafferty, J., & Perez, M. (2016). *Automated vehicle crash rate comparison using naturalistic data*. Blacksburg, VA: Virginia Tech Transportation Research Institute.

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., . . . Morgan, J. F. (2015). *Human factors evaluation of Level 2 and 3 automated driving concepts*. Blacksburg, VA: Virginia Tech Transportation Resarch Institute.

Bolstad, C. A., Cuevas, H. M., Wang-Costello, J., Endsley, M. R., & Angell, L. S. (2008). Measurement of situation awareness for automobile technologies of the future. In G. Rupp (Ed.), *Performance metrics for assessing driver distraction: The quest for improved road safety* (pp. 195–213). Warrendale, PA: SAE International.

Carsten, O., Lai, F. C. H., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control task substitution in semiautomated driving: Does it matter what aspects are automated? *Human Factors*, *54*(5), 747–761.

de Winter, J. C., Happee, R., Martens, M. H., & Stanton, N. A. (2014). Effects of adaptive cruise control and highly automated driving on workload and situation awareness: A review of empirical evidence. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*, 196–217.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*(1), 32–64.

Endsley, M. R. (2017). From here to autonomy: Lessons learned from human-automation research. *Human Factors*. Advance online publication. doi:10.1177/0018720817690634

Endsley, M. R., & Jones, D. G. (2012). *Designing for situation awareness: An approach to human-centered design* (2nd ed.). London, UK: Taylor & Francis.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, *37*(2), 381–394.

Gitlin, J. M. (2016). Another driver says Tesla's autopilot failed to brake; Tesla says otherwise. *ARSTechnica*. Retrieved from http://arstechnica.com/cars/2016/05/another-driver-says-teslas-autopilot-failed-to-brake-tesla-says-otherwise/

Gugerty, L. J. (1997). Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. *Journal of Experimental Psychology: Applied*, *3*, 42–66.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y. C., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors*, *53*(5), 517–527.

Hennan, A., Herdman, C. M., Brown, M. S., & Robert, N. (2014). Effects of coversation on situation awareness and working memory in simulated driving. *Human Factors*, *56*(6), 1077–1092.

Hergeth, S., Lorenz, L., Vilimek, R., & Krems, J. F. (2016). Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving. *Human Factors*, *58*(3), 509–519.

Horswill, M. S., & McKenna, F. P. (2004). Drivers hazard perception ability: Situation awareness on the road. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory, measurement and application* (pp. 155–174). Aldershot, UK: Ashgate.

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomic Science*, *5*(2), 113–153.

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

Ma, R., & Kaber, D. (2005). Situation awareness and workload in driving while using adaptive cruise control and a cell phone. *International Journal of Industrial Ergonomics*, *35*(10), 939–953.

Manzey, D., Reichenbach, J., & Onnasch, L. (2012). Human performance consequences of automated decision aids: The impact of degree of automation and system experience. *Journal of Cognitive Engineering and Decision Making*, *6*, 57–87.

Matthews, M. L., Bryant, D. J., Webb, R. D., & Harbluk, J. L. (2001). Model of situation awareness and driving. *Transportation Research Record*, *1779*, 26–32.

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. (2012). Highly automated driving, secondary task performance, and driver state. *Human Factors*, *54*(5), 762–771.

Metzger, U., & Parasuraman, R. (2001). Automation-related "complacency": Theory, empirical data and design implications. In *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting* (pp. 463–467). Santa Monica, CA.

Neubauer, C., Matthews, G., Langheim, L., & Saxby, D. (2012). Fatigue and voluntary utilization of automation in simulated driving. *Human Factors*, *54*, 734–746.

Petermeijer, S. M., Abbink, D. A., & de Winter, J. C. (2015). Should drivers by operating with an automation-free bandwidth? Evaluating haptic steering support systems with different levels of authority. *Human Factors*, *57*(1), 5–20.

Sabey, B. E., & Staughton, G. C. (1975). Interacting roles of road environment, vehicle and road user in accidents. In *Proceedings of the Paper presented to the Fifth International Conference of the International Association of Accident and Traffic Medicine* (pp. 1–5). London, UK.

Sethumadhavan, A. (2009). Effects of automation types on air traffic controller situation awareness and performance. In *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* (pp. 1–5). Santa Monica, CA.

Soloman, B. (2016, June 30). Tesla autopilot enthusiast killed in first self-driving car death. *Forbes*. Retrieved from http://www.forbes.com/sites/briansolomon/2016/06/30/the-first-self-driving-car-death-launches-tesla-investigation/#69b20eb762d0

Strayer, D. L., & Cooper, J. M. (2015). Driven to distraction. *Human Factors*, *57*(8), 1343–1347.

Treat, J. R., Tumbas, N. S., McDonald, S. T., Shinar, D., Hume, R. D., Mayer, R. E., . . . Catellan, N. J. (1979). *Tri-level study of the causes of traffic accidents: Final report. Volume 1: Causal factor tabulations and assessments* (Publication No. DOT HS-805). Bloomington, IN: Institute for Research in Public Safety.

U.S. Air Force. (2015). *Autonomous horizons*. Washington, DC: U.S. Air Force Office of the Chief Scientist.

Ward, N. J. (2000). Automation of task processed: An example of intelligent transportation systems. *Human Factors and Ergonomics in Manufacturing*, *10*(4), 395–408.

Wickens, C. D., & Kessel, C. (1979). The effect of participatory mode and task workload on the detection of dynamic system failures. *IEEE Transactions on Systems, Man and Cybernetics, SMC*, *9*(1), 24–34.

Young, L. R. A. (1969). On adaptive manual control. *Ergonomics*, *12*(4), 635–657.

Mica R. Endsley is president of SA Technologies and is the former chief scientist for the U.S. Air Force. She received a PhD in industrial and systems engineering from the University of Southern California in 1990. She has published extensively on situation awareness and automation.