ROLE OF SYSTEM STATUS INFORMATION IN THE DEVELOPMENT OF TRUST AND MENTAL MODEL IN AUTOMATED DRIVING SYSTEMS

Michael Manser
Texas A&M Transportation Institute

John Campbell
Exponent

Thomas Fincannon
National Highway Traffic Safety Administration
United States

Audra Krake
Exponent

Liberty Hoekstra-Atwood
Exponent

Caroline Crump
Exponent

Lingtao Wu
Texas A&M Transportation Institute

Paper Number 23-0342

ABSTRACT

Research Question/Objective: In transportation, mental models are essential to mobility and safety because drivers rely on them to understand how to interact properly with their vehicles, the transportation infrastructure, and the environment. Poor performance and errors can occur when a driver acts in accordance with inaccurate mental models. Mismatches between mental models and actual experiences can also lead to reduced trust when, for example, the system with which they interact fails to perform to their expectations. The current study examined differing information types regarding Automated Driving System (ADS) capabilities and limitations on development of mental models and trust while using simulated Level 3 (L3) systems and a "dual model" use case of Level 4 (L4) systems (i.e., the vehicle can be both manually operated and can be controlled by ADS in certain ODDs).

Method and Data Sources: 48 females and males between the ages of 25 and 65 had four exposures to L3 and L4 systems in a driving simulator. Participants used either a basic human machine interface (HMI) that indicated the ADS was active, or they used an enhanced HMI that provided additional information indicating when the system was experiencing limitations (e.g., regarding detection of degraded lane lines). Participants used a simulated Level 3 system for two exposures and a simulated Level 4 system for two exposures. The acquisition and development of mental models and trust were assessed with standardized questionnaires.

Results: Regardless of exposure to each system over time, participants' mental models were more accurate for the simulated Level 4 system compared to the simulated Level 3 system and trust was greater for the simulated Level 4 system during the second exposure.

Discussion and Limitations: This paper summarizes research an ongoing project, and a final report will be published at a later date. Results of the current work suggests that the acquisition and development of mental models and trust can be differentially impacted by how well the ADS performs and the level of automation. However, because the study relied on simulated Level 3 and Level 4 systems, the results may not represent real world implementations of the technology.

INTRODUCTION

At their maturity, Automated Driving Systems (ADS) hold the potential to greatly decrease the number of crashes and save lives. However, there are many important and unanswered questions regarding Level 3 (L3) and Level 4 (L4) ADS [6], particularly around mental models. Mental models refer to a user's knowledge of an automated system's purpose, how it functions, and how it is likely to function in the future [1]. It is therefore important to consider the protentional relationship between a user's mental model of a system and safety. This may be particularly important for L3 vehicles that "cannot guarantee automated achievement of minimal risk condition in all cases within its ODD" and therefore, relies on a fallback-ready user [6]. As described in Campbell et. al. [2], users of L3 vehicles with a functionally accurate mental model are more likely to avoid errors based on incorrect assumptions about system operation and to use the automation appropriately. In contrast to L3, SAE discussion of L4 vehicles states that the system "must be capable of performing the DDT fallback and achieving a minimal risk condition," but also states that these systems "may allow a user to perform the DDT fallback, when circumstances allow this to be done safely" [6], so it may still be important to understand how a user's mental model and trust factors into operation of L4 vehicles. While some drivers may have existing mental models for common automation features such as cruise control, they will have vague or non-existent mental models for early implementations of L3 and L4 ADS [2].

The link between mental models and safety is mediated by trust. Specifically, a person's mental model of an ADS includes their understanding of what a system can and cannot do and it will influence their trust in what the system will do under specific conditions. Research in a variety of domains has identified that a functionally accurate understanding of automated systems is a central aspect to improving users' level of trust of the system [3, 4], where better understanding of the ADS should increase the likelihood that users will have the appropriate level of trust. Inaccurate mental models can lead to both under trust and over trust in an ADS. Research on trust in automation has shown that if a system is unreliable or causes a user to lose trust it will be underutilized and thus, not able to be effective [5]. While L3 and L4 ADS are not currently on the road, it is important to provide an early consideration of trust and its relationship with mental models.

Establishing appropriate levels of trust through functionally accurate mental models is a primary topic of concern in the development and deployment of ADS. When users do not fully understand the system, a mismatch between user expectations and vehicle actions may have a detrimental impact on trust, so it is important to consider how functionally accurate mental models are supported and shaped by an ADS's Human-Machine Interface (HMI). However, despite these established notions, there remain two critical areas in need of examination relative to the relationship between mental models and trust that can inform the HMI design of ADSs. The first area relates to the development of mental models and trust over time when using L3 and L4 ADSs. It is expected that people will begin using L3 and L4 ADSs with mental models and levels of trust based on prior knowledge, likely partly informed through media advertising, news reports, and discussions with peers. However, the continued development of mental models and trust would likely occur through direct interactions with ADSs. This study examines how HMIs can impact mental model and trust development over time.

The second area relates to the HMI implementation of L3 and L4 ADS features. Currently available implementations of HMI provide simple status information to users, typically using very simple binary information (e.g., telltales), such as system ready for activation yes/no and system activated yes/no. However, a relatively simple HMI for ADS may bely the complex nature of its operation and capabilities. It is easy to appreciate that in this case, mental models and subsequent trust in the ADS will be developed based on a limited perspective of the ADS which may not accurately reflect the true nature of the technology. A critical area to examine is how mental models and trust in ADS may be developed through the provision of richer system information that reflects a deeper understanding of the ADS. Specifically, will richer information result in improved/appropriate mental models and levels of trust or will this simply serve to overwhelm or distract users, thus having a negative impact on mental models and trust development. To address this issue, the second goal of the study was to explore the possible benefits of providing information to users about ADS limitations in addition to providing basic status information. In light of these two critical areas, the goals of the study were 1) to assess the impact of varied simulated HMIs on a user's ability to update mental models and appropriately trust L3 and L4 systems over time and 2) assess the possible benefits of providing information to users about the capabilities and limitations on an ADS in addition to providing basic status information.

METHOD

Participants

There were 48 participants in this study. 24 participants were females between 25 and 65 years of age (*M* = 47.8 years, *SD* = 12.6 years) and 24 were males between 25 and 65 years of age (*M* = 43.7 years, *SD* = 12.7 years). To control for potential differences due to demographics, for each HMI treatment level (i.e., Basic versus Information) our goals were to recruit participants so that the group means for age and years of driving experience for females and males would be approximately equal (Basic HMI: females *M* = 44.3 years, *SD* = 14.7 years, males *M* = 46.4 years, *SD* = 11.4 years; Informational HMI: females *M* = 51.3 years, *SD* = 9.4 years, males *M* = 40.9 years, *SD* = 13.9 years; Basic HMI: females mean driving experience of 27.5 years, *SD* = 15.8 years, males mean driving experience of 29.8 years, *SD* =12.2 years; Informational HMI: females mean driving experience of 35.3 years, *SD* = 9.44, males mean driving experience of 24.1 years, *SD* = 14.3 years). The actual mean differences in age and years of driving experience for females versus males were greater than we had hoped and could have affected the results. All participants possessed a valid United States (State of Texas) driver's license, self-reported normal (20/40) or corrected to normal visual acuity, and no color vision deficiencies which may have affected recognition of vehicle-based system icons or human-machine interface elements. To avoid possible experience bias, participants did not have any prior experience with L2 advanced driver assistance systems (ADAS) or L3 ADS technologies.

Apparatus

ADS The ADS was designed to be consistent with SAE J3016 representation of features and operational characteristics of specific implementations of highway automation systems that could be either L3 or L4 ADS [6]. Both the simulated L3 and L4 systems could operate on a four-lane highway with a median in good to moderate weather and could be activated when the vehicle was in "drive" and traveling at least 40 mph. Each simulated L3 and L4 system employed the functional equivalent of: (1) an adaptive cruise control system that would default to 45 mph or a 2s time-headway when a lead vehicle was present and (2) a lane centering system, while also performing the complete object and event detection and response (OEDR) [6]. It is important to note that this study used a short form term L4 to specially mean "dual-mode L4," where the vehicle could be both manually operated and controlled by ADS in certain ODDs. The systems were engaged by pressing a single button on the steering wheel while disengagement could occur when the same button was pressed, the brake or accelerator pedal was pressed, or the steering wheel was turned left or right more than five degrees. The visual HMI icon was positioned between the speedometer and tachometer and depicted a lead vehicle, lane lines, and the distance headway setting to a lead vehicle (see Figure 1).

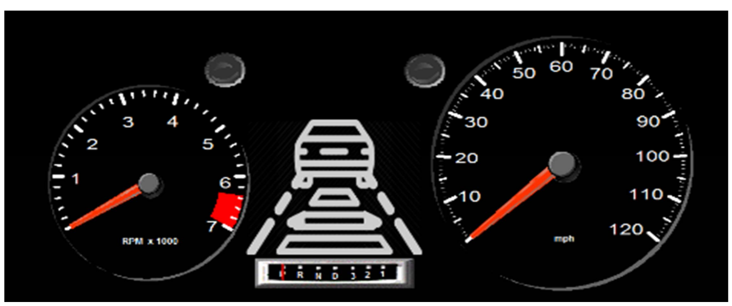

Figure 1. Instrument panel with ADS information available in the center.

Within the Basic HMI condition, when the system became available, the system visual HMI icon appeared white to indicate the system was in standby mode and then, after activation, became green. The icon appeared continuously to inform users of the system status. Within the Informational HMI condition, the system presented the same information as the Basic HMI with the exception that the icon would change from green to yellow, flash at a 2 hz rate, and be accompanied by a two-beep tone when the roadway elements did not provide complete or clear information for the ADS to detect and use. This "limitation" message was triggered due to limitation scenarios (e.g., faded or degraded lane lines, a motorcycle as a lead vehicle, as examples) and the icon would remain yellow once the message was dismissed until the condition that triggered it ended. There was no response or action required from the participants when the limitation message was presented; a key research question was whether or not such information affected their understanding (mental model) or trust in the simulated ADS. In all of the drive segments that included a limitation condition and message, vehicle behaviors in response to the limitation conditions varied depending on whether the L3 or L4 feature was active. Essentially, with the L3 feature active, the vehicle behavior was slightly and temporarily unstable; however, with the L4 feature active, there was no changes in vehicle behavior. For example, in the *Degraded Lane markings in right curve segment*, the L3 ADS exhibited lateral instability in the lane until the lane markings returned to normal, at which time the vehicle returned to stable and accurate lateral lane tracking. For the L4 ADS condition, the vehicle steered to stay in the curve and on the roadway throughout the duration of the limitation condition.

Driving Simulator and Driving World Data collection was conducted in the Texas A&M Transportation Institute's driving environment simulator (manufactured by Realtime Technologies, Inc.) which featured an original equipment manufacturers driver's seat, steering wheel, and accelerator and brake pedals. The visual display consisted of three high resolution monitors providing approximately 160º horizontal and 40º vertical fields of view. Road and ambient noises were provided through a multi-speaker audio system. The driving world simulated a typical highway environment that consisted of a four-lane divided highway with a grass median, Manual on Uniform Traffic Control Devices (MUTCD) compliant roadway markings, and 12 to 15 buildings placed to the right side of each road per mile. Vehicles traveled along each road to mimic light traffic conditions. The roadway environment and traffic were selected to mimic a real-life driving experience in rural areas in which vehicles with L3 or L4 highway automation features would be expected to operate. The driving world was approximately 11 miles in length and required approximately 15 minutes to drive at the posted speed limit of 45 mph.

Drives, Segments, and Scenarios Participants performed four counterbalanced drives (i.e., Drives A1, A2, B1, B2). Each drive consisted of eleven segments with each segment being approximately 1.1 mile long. The first segment, *Start*, allowed participants to accelerate to the posted speed limit and transition from manual to ADS control after being prompted. The final segment, *End*, allowed participants to transition from ADS control to manual driving to exit the highway as the vehicle left its operational design domain. There were four ADS "Normal Driving" segments that contained one scenario each and four ADS "Limitation Message" segments that contained one limitation scenario each. The penultimate segment contained no scenarios for L3 and for L4 drives with the exception that it contained a system automation failure scenario for Drive 2 of L3 only. The penultimate segment allowed for an examination of user responses to an L3 ADS failure. This approach was chosen to demonstrate higher functionality of L4 vehicles. The results of this examination will be presented in future publications. The segment/scenario order for each of the four drives is presented in Table 1, while general segment descriptions are provided in Table 2. It is noted the segments that provided participants with a limitation scenario and resulting limitation message are italicized and underlined, while all other segments were considered "normal driving" in which the ADS did not encounter any limitations within that scenario. The scenario descriptions in Table 2 also summarize the vehicle behaviors that distinguished L3 versus L4 functionality during a limitation condition.

Table 1: Order of Segments/Scenarios within each Drive.

A1 Segment Order	A2 Segment Order	B1 Segment Order	B2 Segment Order
1. Start	1. Start	1. Start	1. Start
2. LVLC	2. RC	2. LC	2. LVPV
3. DMLC	*3. GL*	*3. LVI*	*3. LVMC*
4. RC	4. LVPV	4. LVLC	4. LC
5. LVMC	*5. LVI*	*5. GL*	*5. DMRC*
6. LVPV	6. LC	6. RC	6. LVLC
7. GL	*7. DMRC*	*7. LVSV*	*7. LVI*
8. LC	8. LVLC	8. LVPV	8. RC
9. LVI	*9. LVSV*	*9. DMLC*	*9. GL*
10. ND or RORC	*10.* ND or RORC	*10.* ND or RORC	*10.* ND or RORC
11. End	11. End	11. End	11. End

Table 2: Segment Descriptions.

Abbreviation	Name	Description
	Start	Allowed participants to accelerate to the posted speed limit and transition from manual to ADS control.
ND	Normal Driving	Normal driving operation with no scenarios once ADS activated
DMLC	Degraded Lane Markings in Left Curve	A 90-degree sweeping left curve where the centerline and side lane markings were partially masked due to simulated dirt. In the Basic HMI condition, the HMI remained green with no audible warning. In the Informational HMI condition, an ADS limitation message was presented as the participant traveled next to the degraded markings. In both HMI conditions, L3 ADS exhibited lateral instability in the lane until the lane markings returned to normal. For the L4 ADS condition only, the vehicle was not unstable, and steered to stay in the curve and on the roadway.
DMRC	Degraded Lane Markings in Right Curve	Identical to DMLC with the exception that the curve swept right.
GL	"Ghost" lanes	A section of roadway where one set of faded lane lines appeared offset from brighter lane markings by approximately six to 12 inches to the right (e.g., new lane markings were applied while old lane markings are still visible). In the Basic HMI condition, the HMI remained green with no audible warning. In the Informational HMI condition, an ADS limitation message was presented when the participants' vehicle traveled next to degraded lane markings. In both HMI conditions, the L3 ADS exhibited lateral instability until the lane markings returned to normal, at which time the vehicle returned to stable and accurate lateral lane tracking. For the L4 ADS condition only, the vehicle followed the brighter set of lane lines and did not exhibit any lateral instability throughout.
LC	Left Curve	A 90-degree sweeping left curve where the centerline and side lane markings were fully visible.
RC	Right Curve	A 90-degree sweeping right curve where the centerline and side lane markings were fully visible.
LVLC	Lead Vehicle Lane Change	Participant traveled along road with surrounding traffic changing lanes ahead of the participant's vehicle and ADS engaged. No ADS limitations encountered.
LVMC	Lead vehicle: Motorcycle	Depicted a motorcycle in the left lane moving to the right lane in front of the participant's vehicle. The Basic HMI remained green with no audible warning. In the Informational HMI condition, an ADS limitation message was presented when the motorcycle's wheels crossed the center line. In both HMI conditions, the L3 ADS began to "tailgate" the motorcycle and continued to drive very closely to the motorcycle while remaining near, though below, the set maximum speed of 45 mph. For the L4 ADS condition, the participant's vehicle adjusted its speed to travel two car lengths behind the motorcycle.
LVSV	Lead vehicle: Small vehicle	Identical to LVMC with the exception that a small vehicle moved in front of the participants vehicle.
LVI	Lead Vehicle: Incursion	Scenario entailed a vehicle that pulled into a driver's travel lane and suddenly slowed. The Basic HMI remained green with no audible warning. In the Informational HMI only, an ADS limitation message was presented when the lead vehicle's passenger side wheels crossed the driving line. In both conditions, the L3 ADS decelerated the participant's vehicle in response to the slower lead vehicle. If a driver did not take over within five seconds, the participant's vehicle continued at 40 mph behind the lead vehicle. For the L4 ADS condition only, the vehicle adjusted speed to travel two car lengths behind the lead

		vehicle.
LVPV	Lead Vehicle: Passenger Vehicle	Participant traveled behind a lead vehicle with ADS engaged. No ADS limitations encountered.
RORC	Run Off Road in Curve	The RORC scenario used the DMLC and DMRC scenario. In the Basic HMI, the HMI remained green with no audible warning. In the Basic HMI, the HMI remained green with no audible warning. In the Informational HMI, an ADS takeover request was presented as the participant passed the first degraded markings and continued for 5 seconds. For the L3 ADS Drive 2 only, the vehicle continued to drive straight as the roadway curved until the driver took over. For all the L4 ADS condition exposures, the vehicle steered to stay in the curve and on the roadway.
	End	On-screen message indicates need to resume driving and bring the vehicle to a stop.

Sign Detection Task Participants were asked to engage in a sign detection task to assess attention allocation between the HMI and external roadway. Services road signs were placed on the right and left-hand sides of the roadway at approximately 30s intervals. Each sign presented logos for gas, food, and beverage, lodging, or attractions (see Figure 2) and were different across the four exposures. Participants were given a specific "target" logo to search for at the beginning of each exposure and they then indicated when they detected the target logo by pressing a button on the steering wheel that corresponded to the side of the roadway that the logo was seen (i.e., right or left). There were 24 signs per exposure with approximately one sign with the target logo and two signs without the target logo per segment. Results of the sign detection task will be reported in future publications.

©Texas A&M Transportation Institute

Figure 2. Example of sign detection task road sign.

Questionnaires

A Mental Model and Trust Questionnaire (MMTQ) was developed to understand changes in different components of mental models and trust over time and in response to the two different HMIs. The mental model items included questions about system operation (e.g., how to turn system on and off), participants' understanding of the limitations of the automation features and operator commands (e.g., "what would be your first command in an automated vehicle, such as that used in this study, in the following situation") with multiple choice answers (e.g., brake, stop, change lanes, no response needed, etc.), and questions about participants' understanding of system behavior and operation (e.g., "how would you expect the vehicle used in the study to respond to the following driving situation") with multiple choice answers corresponding to appropriate vehicle behavior (e.g., with L1, L2, L3, or L4/L5 automation). To measure trust, participants were presented with a series of statements (e.g., "Highly automated vehicles can handle unexpected roadway situations" or "Highly automated vehicles are generally safer than human-operated vehicles") and asked to indicate their level of agreement with each statement using a 7-point Likert-scale.) All surveys were administered via SurveyMonkey.

Procedures

Human Subjects Consent, Instructions, Practice Participants read and then signed the human subjects consent form and answered preliminary questions. Participants then completed the MMTQ to collect baseline measures and a simulator sickness questionnaire to screen participants who had a greater propensity to get ill when exposed to the simulator scene. All participants receive the same training and instruction prior to entering the driving simulator. A training video reviewed the interface and operation of the simulated L3 and L4 ADS. Participants then experienced a practice exposure (no ADS and no sign detection task) to become familiar with the operational characteristics of the driving simulator, the task of driving, and to identify participants who exhibited signs of simulator sickness. Participants completed the sign detection task for three minutes while seated in the driving environment simulator (but not driving) to become familiar with the task, completed the second presentation of the MMTQ to assess changes in mental models and trust after instruction, and completed the first administration of the HMI questionnaire.

Drives and Debriefing Participants then completed Exposures 1 through 4 and took, after each drive, the MMTQ and then the HMI questionnaires. The MMTQ administrations after each experimental drive provided an indication of how mental models and trust further matured due to experience with the ADS. The HMI questionnaire administrations after each experimental drive provided the opportunity to assess each interface's usability and effectiveness on the general utility of the ADS information. The order of HMI condition levels (i.e., Basic, Informational) and automation levels (i.e., L3 ADS, L4 ADS) were counterbalanced across exposures (see Table 3). Participants completed an exit questionnaire and received a debrief of the study The study lasted approximately two hours for each participant.

Table 3: HMI and Exposure Counterbalancing by Participant.

Participant Numbers	HMI Group	Order of ADS Levels	Order of Driving Segments
1-3	1 (Basic)	L3, L4	A1, A2 with ROR, B1, B2
4-6	1 (Basic)	L3, L4	B1, B2 with ROR, A1, A2
7-9	1 (Basic)	L3, L4	A2, A1 with ROR, B2, B1
11-12	1 (Basic)	L3, L4	B2, B1 with ROR, A2, A1
13-15	1 (Basic)	L4, L3	A1, A2, B1, B2 with ROR
16-18	1 (Basic)	L4, L3	B1, B2, A1, A2 with ROR
19-21	1 (Basic)	L4, L3	A2, A1, B2, B1 with ROR
22-24	1 (Basic)	L4, L3	B2, B1, A2, A1 with ROR
25-27	2 (Informational)	L3, L4	A1, A2 with ROR, B1, B2
28-30	2 (Informational)	L3, L4	B1, B2 with ROR, A1, A2
31-33	2 (Informational)	L3, L4	A2, A1 with ROR, B2, B1
34-36	2 (Informational)	L3, L4	B2, B1 with ROR, A2, A1
37-39	2 (Informational)	L4, L3	A1, A2, B1, B2 with ROR
40-42	2 (Informational)	L4, L3	B1, B2, A1, A2 with ROR
43-45	2 (Informational)	L4, L3	A2, A1, B2, B1 with ROR
46-48	2 (Informational)	L4, L3	B2, B1, A2, A1 with ROR

Independent Variables

Two types of an HMI were tested that included a "Basic" HMI focusing on on/off status and an "Information" HMI that provided a limitation message, reflecting situations where the system was uncertain about some aspect of lateral or longitudinal control. Two levels of automation were tested, a simulated L3 system and a more capable and better performing simulated L4 system, referred hereafter as "ADS Level". The HMI message served as a limitation message (i.e., for L3 only, this corresponds to a request to intervene) for the conditional driving automation associated with L3 ADS (see page 31 of J3016 [6] for a more detailed description of the conditional nature of L3 ADS) and leaving the decision to respond up to the user. Under L4 operation, there was a limitation message that served as a notification only, not as a request to intervene or a need for fallback performance.

Statistical Approach
The experimenters scored answers to the target questions for each part of the MMTQ separately (i.e., questions about: System Knowledge, System Capabilities/Limitations, and perceived Automation level). The experimenters also created a composite mental models accuracy score (out of 100%) to evaluate overall effects. The composite mental model's accuracy score was derived from 11 questions: 3 system use questions, 2 vehicle behavior questions, and 6 operator command questions.

To test the differences between HMI evaluations and trust ratings (MMTQ) between conditions, the participant's mean ratings for each survey were used as a response variable. The HMI evaluation survey had 5 items and the trust survey had 13 items. In the initial models, these scores were treated as interval data, since the scales were made up of over four Likert-type items that are combined into a composite score [7]. It is important to acknowledge that the Likert scale results are limited: they do not allow for further inferences about the differences in the underlying characteristics reflected in these values (e.g., the meaning of a 0.37 difference in the Trust score).

The data were subjected to linear mixed models. The independent variables used in the models that included all survey administrations were HMI (Basic vs. Information) and Survey Administration (Baseline, Post-Instruction, After Exposure 1, After Exposure 2, After Exposure 3, After Exposure 4). The independent variables used in the models that only included post-exposure survey administrations were HMI, ADS Level (L3 vs. L4), and ADS Exposure (First vs. Second), that is whether the exposure was the first or second time the participant had encountered that ADS level. Participant was treated as a random factor in all models. Non-significant interactions were removed from the final models. The data were analyzed using linear mixed models, built using the lme4 package in R version 4.1.3 (2022-03-10). If the data in a particular model failed to meet all normality assumptions, an alternative model was used. Significant interaction effects were examined with planned post-hoc contrasts using R's multcomp package. Significance levels were set at $p < .05$ where statistical analyses were performed.

RESULTS

Stage 1 Analysis
The results are presented according to the two stages of model testing. The first stage entailed examining what response variables were best predicted by either HMI or Survey Administration (i.e., Baseline, Post-Instruction, After Exposure 1, After Exposure 2, After Exposure 3, After Exposure 4). Results are summarized in Table 4 and indicated that accuracy of mental models questions (i.e., percent questions correct) average trust rating (i.e., average "agreement" scores on the 7-point Likert scale) were each predicted by Survey Administration and that HMI was not a predictor of any response variables. The mental model accuracy analysis significant effect for Survey Administration (95% CI: 11.8, 22.7) indicated that participants scored 17.2% higher on post-training (M = 60.7, SD = 20.3) compared to baseline (M = 43.5, SD = 14.2). The mental model accuracy for system use significant effect for Survey Administration (95% CI: 27.4, 40.7) indicated that participants scored 34.0% higher on post-training (M = 64.9 , SD = 30.4) compared to baseline (M = 30.4, SD = 25.3). The trust analysis significant effect for Survey Administration indicated that trust ratings (-0.32; 95% CI: -0.47, -0.17) decreased between the first (baseline) administration (M = 3.73 , SD = 0.74) and last (exposure 4) administration (M = 3.41 , SD = 0.71).

Table 4: F-Statistics for Final Models Using All MMTQ Survey Administrations

Response Variables	**HMI Condition F-value; *p***	**Survey Administration F-value; *p***
Mental Model Accuracy	$F(1, 46) = 0.007; p = .93$	$F(5, 235) = 21.4; p < .0001$ (17.2; 95% CI = 11.8, 22.7)
Mental Model Accuracy: System Use	$F(1, 46) = 0.07; p = .79$	$F(5, 235) = 46.3; p < .0001$ (34.0; 95% CI: 27.4, 40.7)
Average Trust Rating	$F(1, 46) = 0.10; p = .75$	$F(5, 235) = 6.42; p < .0001$ (-0.32; 95% CI: -0.47, -0.17)

Stage 2 Analysis

The purpose of the second stage was to conduct follow-up linear mixed models contrast analyses to better identify what response variables are best predicted by HMI, ADS Level, and ADS Experience or an interaction between ADS Level and ADS Experience. Table 5 summarizes the results of the contrast analyses. Results indicated that mental model accuracy, mental model accuracy-operator commands, and average trust were each predicted by ADS Level. Results further indicated that each of the response variables were also predicted by the interaction between ADS Level and ADS Experience. Due to the accepted practice that interactions take priority over main effects, the remainder of this discussion will focus on the significant interactions.

Table 5: Results of the Final Models Using All Post-Exposure Survey Administrations.

Response Variable	**HMI Condition F-value; *p***	**ADS Level F-value; *p***	**ADS Experience F-value; *p***	**ADS Level by ADS Experience Interaction F-value; *p***
Mental Model Accuracy	F(1, 46) = 0.003; *p* = .96	F(1, 141) = 47.5; *p* < .0001	F(1, 141) = 0.03; *p* = .86	F(1, 141) = 8.83; *p* < .01
Mental Model Accuracy – Operator Command	F(1, 46) = 0.06 ; *p* = .81	F(1, 141) = 29.6 ; *p* < .0001	F(1, 141) = 0.01; *p* = .94	F(1, 141) = 8.93; *p* < .01
Average Trust	F(1, 46) = 0.002; *p* = .97	F(1, 141) = 23.2; *p* < .0001	F(1, 141) = 2.36; *p* = .13	F(1, 141) = 10.8; *p* <.01

The interaction between ADS Level and ADS Experience for mental model accuracy indicated that the average mental model accuracy scores for L3 and L4 diverged from the first to the second exposure (see Figure 3). There were no significant differences between mental model accuracy between the first and second L3 exposures and between the first and second L4 exposures. However, it is noted that participants exhibited more accurate mental models in the L4 exposures than in the L3 exposures for both the first (L4: M = 69.3, SD = 21.5; L3: M = 63.1, SD = 18.9) (6.25% higher for L4; 95% CI: 0.72, 11.8; *p* < .05) and second (L4: M = 73.8, SD = 20.2; L3: M = 58.0, SD = 19.4) (15.7% higher for L4; 95% CI: 10.2, 21.2; *p* < .001) survey administrations.

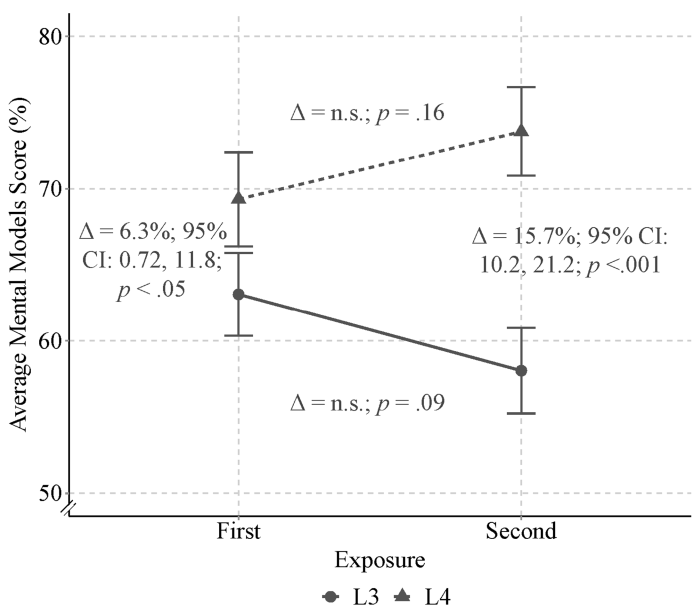

Figure 3. Depiction of the interaction between ADS Level and ADS Exposure on mental model accuracy. The error bars represent standard error.

Results indicated a significant interaction between ADS Level and ADS Exposure for the mental model accuracy –

operator command questions. The interaction was due to more accurate mental models for operator command in the second L4 exposure (M = 72.2, SD = 25.3) compared to the second L3 exposure (M = 53.1, SD = 25.6). Participants' accuracy was 19.1% greater (95% CI: 11.2, 27.0; $p < .001$) after the second exposure to L4 automation than after the second exposure to L3 automation (see Figure 4). There were no significant differences between the first and second L3 exposures, between the first and second L4 exposures, and between the first L4 exposure and first L3 exposure.

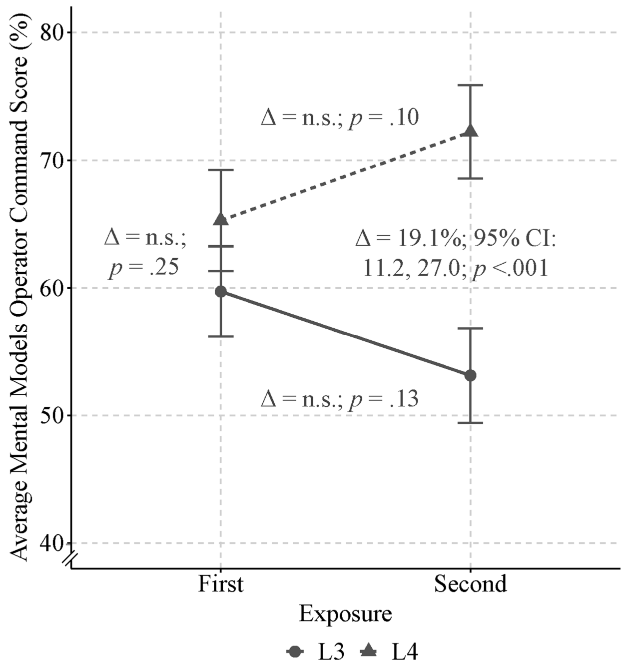

Figure 4. Depiction of the interaction between ADS Level and Survey Administration for mental model accuracy for the operator command questions. The error bars represent standard error.

The analysis indicated that average trust was best predicted by an interaction between ADS Level and ADS Exposure. Trust scores decreased significantly between the first and second L3 exposures (-0.22; 95% CI: -0.38, -0.06; $p < .01$) but not between the first and second L4 exposures. Trust scores were significantly higher after the second L4 exposure (M = 3.65, SD = 0.68) compared to the second L3 exposure (M = 3.28, SD = 0.74) (0.37; 95% CI: 0.21, 0.53; $p < .001$). There was no significant difference between the initial L3 and L4 exposure trust scores.

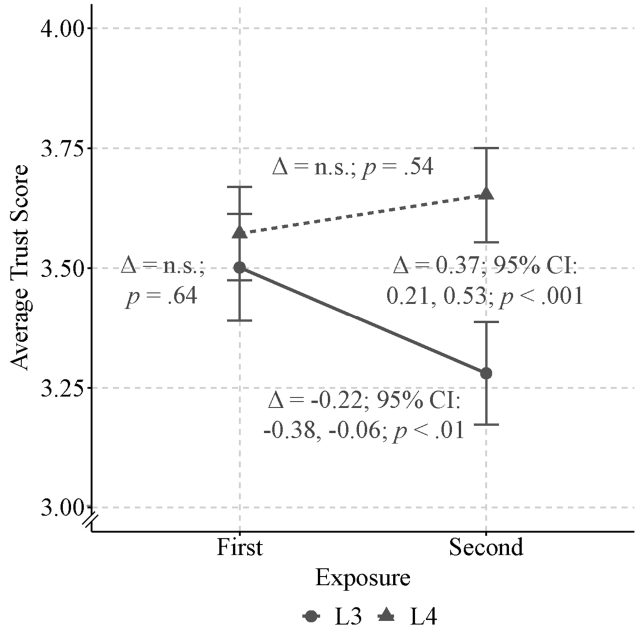

Figure 5. Depiction of the interaction between ADS Level and Survey Administration for trust. The error bars represent standard error.

DISCUSSION/CONCLUSIONS

Mental Models

Previous research indicated that an occupant's mental model of an ADS includes their understanding of what an ADS can and cannot do and will influence their trust in what the system will do under specific conditions. This notion places a critical emphasis on the importance of mental models because they can have a significant impact on not only occupant understanding of ADS but also on the establishment and management of trust over time. The current study sought to assess the impact of HMIs on an occupant's ability to update mental models for both L3 and L4 ADS over time and sought to assess the possible benefits of providing information about the limitations on an ADS in addition to providing basic status information on mental models.

Results of the current study indicated that, in general, mental model accuracy can be impacted by ADS competency and by continued ADS exposure. Specifically, mental model accuracy was generally higher for a better performing simulated L4 system compared to a more limited simulated L3 ADS and that there was a trend for mental model accuracy to increase with greater exposure to the L4 system and to decrease with greater exposure to the L3 system. This same pattern of findings was observed when mental model accuracy relative to operator commands was examined. This pattern of findings suggests that improved ADS understanding was associated with more exposure and with better stability and performance in the L4 compared to the L3 ADS.

It is noteworthy that overall mental model accuracy and mental model accuracy relative to operator commands were not predicted by HMI. Specifically, no differences were found regardless of whether the HMI provided basic "on/off" information or whether the HMI provided information relative to its uncertainty in detecting a potentially hazardous scenario. There were some advantages to the HMI providing limitation information relative at least one scenario (i.e., motorcycle scenario) but the findings were not pervasive across all scenarios. This finding may suggest that additional information about ADS operation may not positively or negatively impact mental model development.

Trust

In light of the notion that trust is mitigated by mental models, it was expected that if there were changes in mental models due to HMI, ADS Level, or ADS exposure, that there would also be changes in trust. Results of the current work indicated that trust indeed changed based on ADS Level and ADS exposure. Specifically, ratings of trust at the first exposure to the simulated L3 or L4 were very similar, but trust in the L3 ADS decreased significantly by the second exposure. In contrast, although not significant, when users were exposure to L4 ADS their trust ratings increased from the first to the second exposure. Overall, the results suggest users' lower understanding of L3 ADS translated into lower levels of trust over time while the better-performing L4 ADS led to slight increases in trust over time.

Limitations

This paper presents an early overview of a study from an ongoing project. Due to the nature of conducting research with ADS vehicles that are not currently on the road, this research effort represents a projection of how these vehicles may operate, and these projections may differ from future implementations of ADS vehicles. For example, future implementations of L3 vehicles may have higher functionality than what appears in this study, and that may impact the development of trust and mental models. Therefore, there may be limitations regarding the degree to which these findings generalize to future research.

The study was presented with several situations that may have served as limitations. First, the study was conducted during the COVID-19 peak which may have impacted the type of people that were willing to participate. The particular concern is that participants willing to engage in socially interactive studies during a pandemic may have inherently different perceptions of trust than participants not willing to engage in studies. The potentially biased participant sample may not accurately reflect responses of the larger population. Second, this study used a group of participants that had limited exposure to ADAS. While this lower level of exposure allowed for the assessment of changes over time, future users may have more relevant exposure to capabilities that provide a better understanding of ADS-equipped vehicles. Third, there were no production-level vehicles available on the market for L3 or L4 features when this study was developed. While the descriptions, characteristics, and behaviors of the L3 and L4 features were consistent with specific implementations discussed in the SAE literature and with the discussions conducted by the research team with industry representatives in a different phase of the overall project, it is certainly the case that real-world implementations of L3 and L4 ADS-equipped vehicle may be different from those implemented or described in this research. For example, L3 vehicles may have better functionality that improves system performance, and these differences may impact the development of mental models and trust. Fourth, mental models and trust are inherently dynamic constructs that can remain stable or change over time and may do so due to a variety of factors. The relatively short study duration may not adequately capture changes in mental models and trust over time. The existence of these limitations suggests the need for future research and caution when extrapolating the results to real world applications.

REFERENCES

1. Goodrich, M. A., & Boer, E. R. (2003). Model-based human-centered task automation: A case study in ACC system design. *IEEE Transactions on Systems, Man, and Cybernetics, Part A, Systems and Humans, 33*(3), 325-336.

2. Campbell, J. L., Brown, J. L., Graving, J. S., Richard, C. M., Lichty, M. G., Bacon, L. P., Morgan, J.F., Li, H., Williams, D.N. and Sanquist, T. (2018). Human factors design guidance for level 2 and level 3 automated driving concepts (Report No. DOT HS 812 555). Washington, DC: National Highway Traffic Safety Administration. https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13494_812555_l2l3automationhfguidance.pdf

3. Lee, J.D., and See, K.A. (2004). Trust in automation; Designing for appropriate reliance. Human Factors, 46(1), pp.50-80.

4. Cassidy, A.M. (2009). Mental models, trust, and reliance: Exploring the effect of human perceptions on automation use (Master's Thesis). Naval Postgraduate School, Monterey, CA.

5. Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 30(3), 286-297.

6. SAE International, (2021). J3016-Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. Warrendale, PA; SAE International.

7. Boone, H. N. J., & Boone, D. A. (2012). Analyzing Likert Data. Journal of Extension, 50(2), 30.