IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 23 May 2023, accepted 25 July 2023, date of publication 4 August 2023, date of current version 9 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3302261

RESEARCH ARTICLE

The Impact of Modality, Technology Suspicion, and NDRT Engagement on the Effectiveness of AV Explanations

QIAONING ZHANG[1], CONNOR ESTERWOOD[1], (Graduate Student Member, IEEE), ANUJ K. PRADHAN[2], DAWN TILBURY[3], (Fellow, IEEE), X. JESSIE YANG[4], AND LIONEL P. ROBERT JR.[1], (Senior Member, IEEE)

[1]School of Information, University of Michigan, Ann Arbor, MI 48109, USA
[2]Department of Mechanical and Industrial Engineering, University of Massachusetts Amherst, Amherst, MA 01003, USA
[3]Robotics Department, University of Michigan, Ann Arbor, MI 48109, USA
[4]Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109, USA

Corresponding author: Qiaoning Zhang (qiaoning@umich.edu)

This work was supported by the University of Michigan (MCity) through the University of Michigan Office of Research.

ABSTRACT Explanations — reasons or justifications for action — are being used to promote the acceptance of automated vehicles (AVs). Yet, it is unclear whether and how the modality of explanation affects its effectiveness. Despite its importance in the technology acceptance literature, the impact of technology suspicion on the adoption of AVs is yet to be fully examined. To expand our understanding of AV explanation, we conducted a within-subjects experiment with 32 participants using a high-fidelity driving simulator. Four conditions were presented to participants: (1) auditory explanation with a non-driving-related task (NDRT), (2) auditory explanation without NDRT, (3) visual explanation with NDRT, and (4) visual explanation without NDRT. The results indicate that auditory explanations are more effective in reducing anxiety and unsafety perception for high-suspicion individuals, especially in the absence of NDRT. Conversely, individuals who are less technology suspicious prefer visual explanations, which can result in lower levels of anxiety and perceived unsafety. The study highlights the importance of considering individuals' technology suspicion and engagement with NDRT when selecting the appropriate explanation modality, and the findings can guide the design of future AV systems to promote effective human-machine interaction.

INDEX TERMS Automated vehicle (AV), explanation, advanced driver assistance system (ADAS), modality, non-driving-related task (NDRT), technology suspicion, anxiety, unsafety.

I. INTRODUCTION

In the effort to promote the acceptance of automated vehicles (AVs) [1], explanations are being used to convey the reasons or justifications for the vehicle's actions. AVs at levels 4 and 5 driving automation, as defined by the Society of Automotive Engineers (SAE), take full responsibility for all driving tasks in some or all circumstances, respectively [2]. For the purposes of this study, AVs refer to SAE levels 4 and higher vehicles. Explanations help increase transparency and assist individuals in forming a correct mental model of their driving environment [3], [4]. By making the AV's actions predictable and understandable to humans [3], [4], [5], explanations are effective in reducing fears, concerns, and driving-related anxiety, ultimately promoting trust [3], [4], [5], [6], [7], [8].

Modality refers to the sensory channel used by AVs to communicate with humans. AV explanations are typically delivered through two uni-modal modalities: auditory and visual. However, little is known about the differential effects of these modalities on explanation effectiveness [1]. This is problematic because previous research has suggested that modality effects can significantly impact communication between level 1-3 vehicles and humans. While the auditory modality has been shown to promote safety by alerting individuals when their vehicle is too close to the vehicle in front of them [9], [10], information delivered

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

through this modality can also be annoying and startling to individuals [11], [12], [13]. Moreover, the effectiveness of auditory cues is moderated by the involvement of non-driving-related tasks (NDRTs), as they can relieve individuals' visual attention burden [14], [15], [16]. On the other hand, the visual modality requires shorter information recognition times and supports continuous information awareness [10], [17]. However, the information provided through the visual modality can also divert attention away from the driving environment, compromising safety [18]. Overall, a better understanding of the effects of different modalities on AV explanations is crucial for designing effective AV explanations that enhance user understanding, engagement, and safety.

An effective AV explanation could depend not only on the explanation itself but also on the individual characteristics of the person receiving the information. One such individual trait that has been linked to attitudes and acceptance of technology is technology suspicion. Technology suspicion is defined as "a person's simultaneous state of cognitive activity, uncertainty, and perceived malintent about underlying information that is being generated, collated, sent, analyzed, or implemented by an external agent" [19, p. 495]. Understanding technology suspicion is critical in comprehending users' attitudes towards technology [20], [21], its adoption [20], [22], [23], and eventual use [21], [24], [25], [26]. Despite its importance, little research has investigated the impact of technology suspicion on the effectiveness of AV explanations.

To expand our understanding of AV explanation effectiveness, this study addressed the following question: Does the modality of the AV explanation impacts its effectiveness, and does technology suspicion and NDRT engagement influence this effect? Building on theories and empirical evidence, we hypothesized that modality effects are moderated by the NDRT involvement and an individual's levels of technology suspicion, both of which are key factors in human-technology collaboration. We conducted a within-subjects experiment with 32 participants using a high-fidelity driving simulator. The experimental design presented each participant with four AV explanation conditions: (1) auditory explanation with the NDRT, (2) auditory explanation without the NDRT, (3) visual explanation with the NDRT, and (4) visual explanation without the NDRT. Our results show that modality is essential in determining the effectiveness of AV explanations and that an individual's suspicion and engagement with NDRT play a role in this effect.

This study makes several key contributions. Firstly, it sheds light on the significance of modality in determining the effectiveness of AV explanations. Secondly, it shows how task demands and individual attributes can considerably influence this impact. Lastly, the study offers design recommendations to enhance the acceptance and effectiveness of AVs.

II. BACKGROUND

A. AUTOMATED VEHICLE EXPLANATIONS

In the realm of automated vehicles (AV), an AV explanation refers to the information provided by the vehicle's system to the users about its behavior or actions [1]. In this study, the term "AV" refers to SAE levels 4 and higher vehicles, where the automated driving system takes full responsibility for all of the driving tasks in some and all circumstances, respectively [27]. The literature exploring the impact of explanation on AV acceptance can be broadly categorized into two areas: AV explanation timing and AV explanation content.

Regarding AV explanation timing, studies show that the effectiveness of AV explanations can be greatly impacted by when the explanation is presented to the individual [1]. Providing Regarding timing, studies show that the effectiveness of AV explanations can be greatly impacted by when the explanation is presented to the individual [1]. Providing the AV explanations before the AV performs the actions leads to positive emotions (e.g., trust and preference) and decreases negative feelings (e.g., anxiety and workload) [3], [4], [5], [7], [28]. By contrast, explanations following actions have led to the lowest levels of AV trust and preference, even though they improved individuals' understanding of the driving system [3], [7]. The impact of timing is also influenced by an individual's age: younger individuals tend to have less-negative reactions to explanations provided after the AV has taken action [29].

The second area examines how the effectiveness of AV explanations can be greatly influenced by the content of the explanation presented to the individual [1]. There are generally three types of explanation content: why-only (i.e., why an AV would take action), what-only (i.e., what the AV would do), and why + what (a combination of why-only and what-only explanations) [1]. Research indicates that why-only explanation content leads to the best outcomes in terms of promoting positive attitudes such as acceptance, trust, preference, understandability, alertness, and a sense of control while decreasing negative feelings such as anxiety and increasing feelings of safety [5], [6], [30]. On the other hand, what-only explanation content has been linked to the worst outcomes, resulting in the most dangerous driving behaviors and the least acceptance of the AV [6]. The why + what explanation is valued for promoting positive emotional valence and safe driving, but individuals felt anxious and annoyed when they received the what + why explanation [4], [6], [31].

To summarize, previous research has contributed to our understanding of AV explanation efficacy by exploring the effects of content and timing, but has not examined whether modality impacts AV explanation effectiveness [1]. Modality refers to the classification of a single independent sensory input/output channel between automation and a human [32], and in the context of AV explanations, it refers to the sensory input/output channels used to convey information to the users about the vehicle's behavior or action. The literature has used two types of modality in AV explanations, auditory and visual, with auditory explanations typically provided in a standard American accent in a neutral tone of voice [3], [4], [5], [6] and visual explanations presented using text and abstraction [30]. However, the impact of modality on the effectiveness of AV explanations has yet to be specifically investigated in prior research [1].

B. COMMUNICATION MODALITY AND DRIVING AUTOMATION

While little is known about the impact of modality on explanations provided by levels 4 and above AVs, previous research has investigated the impact of alert modality for levels 1-3 of driving automation. Research on driving-related information display has mainly focused on two uni-modal modes of communication: visual and auditory. Visual modalities use a variety of cues, such as color, pattern, and text, to convey information [33]. Auditory modalities use sound to transmit information to a user interacting with a system [34]. These modalities are typically designed to provide individuals with time-critical warnings (e.g., beeping or icons), commands (e.g., takeover requests), and contextual information (e.g., road sign highlights). However, displaying information visually or stating it verbally has both advantages and disadvantages for individuals. In the following section, we will discuss this research area in more detail.

1) AUDITORY MODALITY AND DRIVING AUTOMATION

Numerous studies investigating driving automation levels 1-3 have shown that the auditory modality is superior to the visual modality [9], [10], [31], [35], [36], [37], [38], [39]. Since driving relies heavily on visual resources, auditory information imposes fewer cognitive demands on the driving task, rendering it more effective than visual warnings, particularly in critical driving situations [10], [36].

Previous research on level 3 vehicles discovered that auditory-only takeover requests resulted in much shorter takeover times than visual signals [31], [37], [38]. Furthermore, research on level 1 or 2 driving automation demonstrated that auditory warnings can benefit drivers [9], [10], [35], [39]. For example, the auditory headway warning alert can enhance safety by assisting drivers in maintaining a safe distance from the car in front of them [10].

This was especially evident when drivers were distracted by non-driving-related tasks (NDRTs). In situations with highly automated systems, drivers can afford to pay less attention to the traffic situation or the system, so they are more likely to engage in NDRTs such as reading, watching videos, playing games, or making phone calls [14], [15], [16]. Studies have suggested that auditory cues are particularly beneficial when drivers are not paying complete attention to the road and are engaged in NDRTs because the auditory cues can reduce the driver's visual attention load and serve as a substitute for vision [40], [41], [42]. Auditory warnings are less distracting and better aid in attention orientation compared to visual warnings [9], [35], [39], making them a more suitable option for providing warnings and quickly communicating the potential danger's magnitude [10].

It explains why individuals are more inclined to rate vehicles presenting information in an auditory format as more satisfactory, useful, effective, and safe [34], [43]. Nevertheless, the auditory modality can cause annoyance and startle individuals. According to researchers, people experienced increased levels of perceived annoyance when presented with auditory rather than visual information [13]. Moreover, auditory information can startle drivers, resulting in increased stress, delayed action, or wrong action [11]. This claim was supported by Gupta et al. [12], who found that participants pressed the accelerator harder after the auditory warning started. Therefore, while the auditory modality has many benefits, it may also have negative consequences in some situations.

2) VISUAL MODALITY AND DRIVING AUTOMATION

Research on driving automation levels 1-3 has also explored the use of visual information. Recent studies on level 3 vehicles suggested that visual alerts, such as visual icons in the head-up display, can enhance perceptions of ease of use, transparency, and satisfaction [44], [45]. In particular, visual alerts were rated higher than auditory cues because they were found to be easier to understand, and background noise had no negative impact on their effectiveness [44].

The benefits of visual alerts in levels 1 and 2 driving automation have also been highlighted in previous research. Visual alerts can help drivers maintain continuous awareness of the surrounding environment by depicting the road environment and alerting them to obstacles in blind spots [10]. Visual warnings, such as texts and icons, require shorter recognition times of warning urgency than auditory warnings [17]. Empirical findings suggest that the duration of a visual message (1.8 seconds for icons, 3.6 seconds for text) is shorter than the time required to perceive and comprehend the same information through an auditory modality (less than 10 words, about 5 seconds) [9], [46].

Prior research has also evaluated the effectiveness of visual alerts in the presence of NDRTs. It was found that visual alerts may be insufficient in directing individuals' attention to their driving tasks, particularly in urgent situations [37], [38], [47], [48].

3) SUMMARY OF MODALITY IMPACTS

Research investigating levels 1, 2, and 3 of driving automation has demonstrated the significant impact of modality (auditory or visual) on the effectiveness of alerts. However, little to no attention has been given to examining how modality influences the effectiveness of explanations, which typically contain longer and more complex content, particularly for levels 4 and above autonomous vehicles (AVs).

C. TECHNOLOGY SUSPICION

Understanding technology suspicion is crucial for understanding users' attitudes towards technology, adoption, and eventual performance. In the field of AVs, technology suspicion refers to the user's state of cognitive activity, uncertainty, and perceived malintent about the information being generated, collated, sent, analyzed, or implemented by the AV system [19]. The literature provides several reasons why technology suspicion is important.

First, technology suspicion negatively correlates with the intention to adopt automation [20]. High suspicion erodes confidence in judgment, thereby reducing the reliance on and adoption of automated tools. For example, consumers initially resisted using ATMs for transactions because of harbored suspicions [22]. Additionally, people are more likely to reject innovations if they are suspicious of data security, privacy protection, and system reliability [23].

Second, prior research has shown that suspicion is crucial for automation performance. Highly reliable automation fosters complacency among users, which can result in catastrophic performance errors and system accidents [24], [25], [26]. Low levels of suspicion lead automation operators to over-rely on and be overconfident in automation, which, in turn, makes them incapable of detecting malfunctions in automation [21], [26].

Finally, suspicion significantly affect attitudes (i.e., anxiety, fear, and trust) in the context of technology. Previous research suggests that the malicious intent and uncertainty aspects of increased suspicion can enhance emotional arousal, resulting in anxiety and fear [19]. Additionally, Muir posited that the stability of trust is determined by the level of suspicion an individual holds [20]. Specifically, if people hold low levels of suspicion, they are less likely to generate alternative explanations for inconsistent behavior and attribute malicious intent to a given entity. This, in turn, affects the stability of individuals' trust [20], [21].

In sum, research has advanced our understanding regarding the effects of explanation (i.e., content and timing), modalities, and technological suspicion. However, the manner in which these factors interact and influence perceptions and adoption of AVs remains unclear.

III. HYPOTHESIS DEVELOPMENT

This study aims to investigate the relationship between explanation modality, non-driving-related task engagement, technology suspicion, and explanation effectiveness in the context of automated vehicles (AVs). The theoretical model used in this study, as depicted in Figure 1, was constructed based on Multiple Resource Theory (MRT), empirical evidence on the relationship between suspicion level and information needs, and bottleneck theory of attention. Specifically, we hypothesized and examined the three-way interaction between these factors and their influence on the effectiveness of AV explanations.

Multiple Resource Theory (MRT) postulates that humans have separate fixed-capability resources for processing information. These resources are determined by four categorical and dichotomous dimensions: processing stages (cognition and perception), perceptual modalities (visual and auditory),

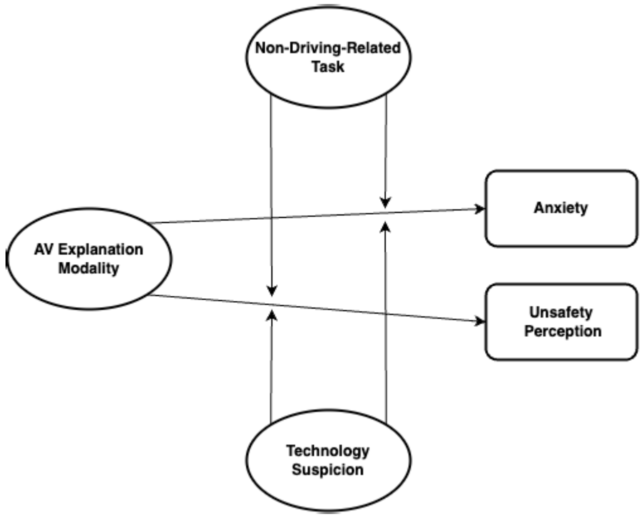

FIGURE 1. The proposed research model depicting the three-way interaction between explanation modalities, NDRT engagement, and technology suspicion in relation to anxiety and unsafety perception.

visual channels (focal and ambient), and processing codes (spatial and verbal) [49]. As individuals have a limited capacity to process information, they tend to pay attention to only a portion of the information in their environment, neglecting the rest [50]. When multiple tasks require the same attentional resources, there is a higher likelihood of producing significant task interference. Therefore, it is best for tasks to draw upon separate resources from distinct perceptual modalities (e.g., auditory-visual) rather than from similar modalities (e.g., visual-visual) [51], [52]. According to the MRT model, auditory explanations are more effective than visual explanations in processing information, as visual attention is required to monitor and comprehend the driving environment. Visual explanations require the same attentional resources, which interferes with the ability of individuals to achieve their goals.

The level of technology suspicion can affect the effectiveness of the explanation modality because people's need for explanation can vary. Research shows that individuals with high technology suspicion are more likely to evaluate the autonomous vehicle's driving behavior and pay more attention to processing the explanation, compared to those with low technology suspicion [20], [21]. High-suspicion individuals value effective explanations more because of their heightened demands to evaluate the AV's performance actively. Auditory explanations can provide these individuals with clear and easy-to-understand justifications for the actions of the vehicle, which can help alleviate negative feelings, such as anxiety and perceived unsafety. On the other hand, low-suspicion individuals may find auditory explanations disturbing and interpret them as a problem with AV driving. Since they are less likely to monitor and evaluate the AV's behavior, they have a low need for the AV to explain its behavior. Although auditory explanations can reduce task interference and enhance comprehension of the driving situation, low-suspicion individuals may not require or appreciate such explanations and could even view them as a warning of a more serious problem [11], [12], [13]. Therefore, using auditory explanations may increase fear and anxiety as well as the perception of being unsafe for low-suspicion individuals.

In cases where individuals are engaged in non-driving-related tasks (NDRTs), the impact of the explanation modality may be altered, regardless of their level of technology suspicion. Bottleneck theory suggests that when two tasks require the mechanism simultaneously, a bottleneck occurs and one or both tasks may be delayed or affected [53]. Because individuals have limited attentional resources, they filter information and stimuli so that only the most salient information is absorbed [54], [55]. In situations where NDRTs are required, individuals may allocate their attentional resources to the task, and the AV explanations are likely to be filtered and not have a significant impact on their feelings of anxiety and unsafety, regardless of the modality. In other words, when individuals are engaged in an NDRT, the differences between the explanation modalities are suppressed.

In summary, the relationship between explanation modality and the effectiveness of AV explanation is dependent on both NDRT engagement and technology suspicion. We hypothesized that offering auditory explanations instead of visual ones helps decrease (a) anxiety and (b) unsafety perception among people with high technology suspicion who are not engaged in an NDRT because this mode of explanation interferes less with environment monitoring and helps redirect their attention. Our theoretical arguments are depicted in Figure 1.

H1: A three-way interaction is expected among **explanation modalities, NDRT engagement, and technology suspicion** regarding anxiety and unsafety perception, such that auditory explanation is more effective than visual explanation in reducing (a) anxiety and (b) unsafety perceptions for high-suspicion individuals not engaging in an NDRT.

IV. METHOD

This study employed a within-subject, three-factor repeated-measures design to evaluate the effects of explanation modality (visual vs. auditory) and non-driving-related task (NDRT) engagement (with NDRT vs. without NDRT) on two dependent variables: anxiety and unsafety perceptions. The research was carried out in compliance with the American Psychological Association code of ethics and was approved by the university's institutional review board. Informed consent was obtained from all participants.

A. PARTICIPANTS

A total of 32 individuals (mean age = 23.6 years; standard deviation = 3.6 years) participated in this study. The recruited participants comprised 12 women and 20 men with an average driving experience of 7.0 years. Before conducting the study, a power analysis was performed to determine the appropriate sample size. Based on Cohen's [56] criteria, the effect size in this study was considered large at 0.8 [56]. Using GPower3.1 with alpha = 0.05 and power = 0.8, the total sample size needed for the "ANOVA: repeated measures, within-between interaction" group comparison was less than 24. Therefore, the 32 participants in this study were more than sufficient to produce statistically significant results. None of the participants had previously participated in any study that addressed similar research questions. Participants were screened for driver's license status, visual impairments, and simulator sickness susceptibility. A $20 compensation was given to each participant.

B. STUDY DESIGN

Four conditions (see Table 1) are developed to test the hypothesis. Two modalities of AV explanation and two levels of NDRT engagement (i.e., videos) are represented by these four conditions: auditory explanation with NDRT (AW), auditory explanation without NDRT (AN), visual explanation with NDRT (VW), and visual explanation without NDRT (VN). The foundation for designing the four AV explanation conditions in this study is based on the three factors that were hypothesized to impact the effectiveness of AV explanations: explanation modality (auditory vs visual), non-driving-related task (NDRT) engagement (present vs absent), and technology suspicion levels (high vs. low). The four conditions were created by crossing the two levels of explanation modality (auditory vs visual) with the two levels of NDRT engagement (present vs absent) while accounting for the technology suspicion level. By manipulating these factors, the study aimed to investigate the impact of AV explanation on anxiety and unsafety perception in a driving scenario. We employed a Latin square design to avoid any potential bias in the order of the four AV explanation conditions.

Within each AV explanation condition, which lasted approximately 6-8 minutes, three unexpected and unique events (e.g., events caused by other drivers, events caused by police vehicles, and events triggered by unexpected reroutes) occurred within three environments (urban, highway, and rural). Each event differed in terms of its surrounding environment and traffic situation. To minimize potential confounding effects arising from sequence and order, we employed a counterbalanced approach (i.e., Latin square design) to ensure an even distribution across the four AV explanation conditions. In accordance with the prescribed procedure, events occurred at intervals of 1-2 minutes.

C. INDEPENDENT VARIABLES

1) EXPLANATION MODALITY

The two types of modality used in this study were auditory and visual. Auditory explanations were delivered through the simulator's audio production system and a cab interior sound system. The AV provide the explanation of what it was going to do 7 seconds prior to taking the actual actions in a neutral tone of voice delivered by a man with a standard American accent. Visual explanations were also provided 7 seconds before the AV took action. However, unlike the

TABLE 1. Experimental Design.

		Non-driving related Task (NDRT) Engagement	
		With NDRT (W)	Without NDRT (N)
AV Explanation	Auditory Explanation (A)	AW	AN
Modality	Visual Explanation (V)	VW	VN

auditory explanations, the visual explanations were displayed on a 7-inch monitor connected to the simulator dashboard with a high-definition resolution of 1,024 × 600. The explanation texts were set to 48 font size, displayed with black characters on a white background. The information regarding the explanation content for each event can be found in Table 3.

2) NON-DRIVING RELATED TASK

For the conditions with an NDRT, participants were given instructions to do activities during the automated driving of the AV simulator. They were required to take part in two question-and-answer activities prior to the second and third events. Each activity involved watching a science popularization video and responding to a question related to the video (NDRT examples link: https://youtu.be/H5BsEXoNZ04). After the researchers instructed participants to "Please start watching the video," participants watched a video using a 10.2-inch iPad placed on the simulator's center console, which was well within reach of the participants. Each video lasted 30-90 seconds and was followed by a video-related question. The activity start and end times were calculated and programmed so that they would not overlap with the AV event. For the purpose of examining how participants would allocate their attention and evaluate the AV differently, the participants were not informed of when the AV event would take place while engaged in the NDRT.

3) TECHNOLOGY SUSPICION

We used the Complacency Potential Rating Scale [26] as a measure of technology suspicion. The technology suspicion scores were dichotomized into two groups based on their means: high or low. Specifically, 15 participants had scores above the mean and were classified as having high suspicion, while 17 participants had scores below the mean and were classified as having low suspicion. The scale consisted of four sub-dimensions relating to complacency: confidence, reliance, trust, and safety, with a total of 12 items. The scale ranged from 1 to 7 (1: strongly disagree; 7: strongly agree).

D. DEPENDENT VARIABLES

The dependent variables included anxiety and unsafety perception. All of the questionnaires employed are shown in Table 2.

1) ANXIETY

We measured anxiety using Nass's 7-point rating scale (1: very poorly; 7: very well), which is a published model from the CHIMe Lab at Stanford University used to measure individual attitude [57]. The measure of anxiety comprised the average responses to four adjective items reflecting feelings while driving the AV: fearful, afraid, anxious, and uneasy.

2) UNSAFETY

Unsafety refers to a combined measure of the probability and consequence of a mishap that could result in a loss event (i.e., fatalities or injuries) [58]. Eight items, taken from Hayes et al. [59], dealt with an individual's perceptions of AV unsafety, including dangerous, safe, hazardous, risky, could get hurt easily, uneasy, the chance of death, and scary. Participants were asked how much they agreed that the item described their perception of unsafety while driving with the AV. The respondents rated from "strongly disagree" to "strongly agree" on a 5-point scale.

E. APPARATUS

The study was conducted in a simulated environment using a high-fidelity advanced driving simulator (Figure 2). Two components made up this simulator: a Nissan Versa sedan providing all manual controls and a simulation system operating with version 2.63 of Realtime Technology's simulation engine SimCreator as well as a programmable software package to control automatic vehicle functions. This virtual driving environment was displayed to the participant on four flat screens. The forward road scenes were projected onto three screens about 16 feet in front of the participant (120-degree field of view), and the rear sight was shown on a back screen located 12 feet from the steering (40-degree field of view). The forward screens were each set at 1,400 × 1,050 pixels and updated at 60 Hz, and the rear screen was set at 1,024 × 768 pixels.

Automated driving features of the driving simulator employed in this study were designed to simulate an AV with SAE level 5, in which the individual was not required to actively monitor the environment. The AV was able to automatically carry out and respond to driving conditions under all roadway and environmental conditions including longitudinal and lateral vehicle control, navigation, and responses to traffic control devices and other elements of the road. To begin a simulated drive, participants were instructed to push a button on the steering wheel. When the automated driving was engaged, participants no longer needed to take control of the vehicle.

The event explanations for four driving conditions were selected from previous literature and corresponded to realistic unexpected situations in automated driving as shown in Table 3 [5], [60], [61], [62], [63], [64].

TABLE 2. Factor Loading.

Variable	Items	Cronbach's Alpha	Component		
			1	2	3
Technology Suspicion	Automated systems used in modern aircraft, such as the automatic landing system, have made air journey safer.	0.75	0.66		
	ATMs provide safeguard against the inappropriate use of an individual's bank account by dishonest people.		0.70		
	Automated devices used in aviation and banking have made work easier for both employees and customers.		0.73		
	Automated devices in medicine save time and money in the diagnosis and treatment of disease.		0.72		
	Bank transactions have become safer with the introduction of computer technology for the direct deposit of checks.		0.68		
Anxiety	Fearful	0.93		0.71	
	Anxious			0.69	
	Afraid			0.79	
	Uneasy			0.76	
Unsafety	Hazardous	0.87			0.84
	Risky				0.87
	Could get hurt easily				0.82
	Unsafe				0.84
	Chance of death				0.60

TABLE 3. Event Description.

Drive #	Event	Explanation
Drive 1	Swerving Vehicle Ahead	"Swerving vehicle ahead, Slowing Down."
	Efficiency Route Change	"Road obstruction ahead, Rerouting."
	Stopped Police Vehicle on Shoulder	"Emergency vehicle on shoulder, Changing lanes."
Drive 2	Heavy Traffic Rerouting	"Traffic reported ahead, Rerouting."
	Oversized Vehicle Ahead	"Oversized vehicle blocking roadway, Slowing down."
	Police Vehicle Approaching	"Emergency vehicle approaching, Stopping."
Drive 3	Abrupt Stopped Truck Ahead	"Road obstruction ahead, Changing lanes."
	Road Hazard Rerouting	"Road obstruction ahead, Rerouting."
	Stopped Police Vehicle on Shoulder	"Emergency vehicle on shoulder, Changing lanes."
Drive 4	Police Vehicle Approaching	"Emergency vehicle approaching, Pulling over and stopping."
	Unclear Lane Markings Rerouting	"Unclear lane lines, Rerouting."
	Vehicle with Flashing Hazard Lights Ahead	"Vehicle with hazard lights ahead, Slowing down."

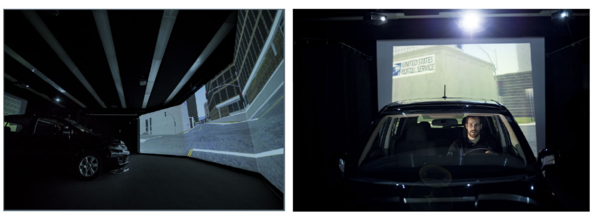

FIGURE 2. Driving Simulator.

F. PROCEDURE

A consent form was provided to participants on arrival at the simulator lab. The form described the experiment and outlined the individual's informed consent. Prior to taking part in the study, all participants also completed a pre-experiment survey on an iPad that collected information such as their age, gender, dominant hand, prior experience with technology (i.e., driving video games, computer, tablet or other touchscreen devices, and driving assistance systems), and technology suspicion.

Participants were given a 3-minute training session prior to the main study in order to familiarize themselves with the driving environment and the simulator. The participants were instructed on the attributes of the AV during the training session and were informed that the vehicle could function safely entirely on its own. In particular, the vehicle was able to function in all driving conditions in the same manner as the average human driver; it complied with all traffic laws; it received navigation information from external sources, such as Google Maps, and could change routes to reach a destination more rapidly if one were identified or available; and it maintained lanes by visually sensing the lane markings on the roadside. Participants were instructed to switch the AV from manual to automatic control by placing the AV in the center of the right lane and pressing the automated mode activation button on the steering wheel. To engage the NDRT, participants were instructed to follow the researchers' signal (i.e., "Please start watching the video") and press the video play button on the iPad.

The 60-minute main study session followed the training. The participants experienced all four experimental conditions as shown in Table 3. Participants took part in a 6- to 8-minute drive with events occurring at intervals of 1-2 minutes at prescribed times in each AV explanation condition. They were also given a question to be answered after each NDRT video ended (e.g., Which one is the biggest contributor to the greenhouse effect? A. Methane; B. CO2; C. Nitrous Oxide). After each of the driving conditions, participants were asked to fill out a follow-up survey consisting of two questionnaires to measure their anxiety and perceptions of unsafety. All questionnaire items were adapted from previously validated research. Participants were given a 2-minute break between AV explanation conditions. The procedure of the study is depicted in Figure 3.

V. RESULT

A. RELIABILITY AND CONSTRUCT VALIDITY

To access construct validity and reliability, we used a sequential approach suggested by Wille [65], [66]. First, using a process that sequentially removes items that are internally inconsistent, the reliability of each construct is enhanced to reach an acceptable threshold of .70. The second step involves assessing and modifying the discriminant validity of the construct by removing any items that are significantly loaded on more than one factor using an exploratory factor analysis (EFA). In addition, convergent validity is increased by identifying and discarding items that fail to load significantly on any of the factors. Table 2 shows the refined constructs and corresponding items, which all met the acceptable reliability threshold of .70 or above and loaded at .60 or above. An overview of the correlations, standard deviations, and means can be found in Table 4.

B. HYPOTHESIS TESTING

We tested our hypothesis on a sample of 32 participants. To examine how explanation modality, NDRT engagement, and technology suspicion affected the dependent variables (i.e., anxiety and perceived unsafety), we performed linear mixed models to find out whether there was a significant difference between the means of three or more independent groups based on these three distinct factors. Subjects were treated as random effects to resolve non-independence in all the linear mixed models. Statistical analysis was performed using IBM SPSS 28.0 statistics software. The alpha was set at 0.05 for all statistical tests. A Bonferroni alpha correction was applied to all post hoc comparisons.

1) THE EFFECT OF EXPLANATION MODALITY, NDRT ENGAGEMENT, AND TECHNOLOGY SUSPICION ON ANXIETY

In Hypothesis 1a, we predicted that there would be a three-way interaction involving explanation modality, NDRT engagement, and technology suspicion with anxiety. As shown in Table 5, the three-way interaction of explanation modality, NDRT engagement, and technology suspicion with anxiety was significant ($F = 4.38$, $p = 0.04$). Thus, H1a was supported. Figure 4 depicts the three-way interactions related to hypothesis 1a. Using simple slopes analysis, we found that explanation modality did not significantly affect anxiety when high-technology-suspicion individuals engaged in the NDRT (Slope 1; $t = -0.07$; $p = 0.95$). The same pattern was evident in those with low suspicion because there were no significant differences in anxiety between auditory and visual explanations when there was an NDRT (Slope 3; $t = 0.54$; $p = 0.59$).

Although auditory explanation was not helpful when the NDRT was engaged, we found that individuals with high technology suspicion were less anxious with auditory explanation than visual explanation when no NDRT was presented (Slope 2; $t = 2.87$; $p = 0.01$). Interestingly, the effect of explanation modality on anxiety was significant in the opposite direction for individuals with low technology suspicion. It was found that low-suspicion individuals felt less anxious when visual explanations were provided instead of auditory ones (Slope 4; $t = -2.09$; $p = 0.05$).

2) THE EFFECT OF EXPLANATION MODALITY, NDRT ENGAGEMENT, AND TECHNOLOGY SUSPICION ON UNSAFETY PERCEPTION

On the perception of unsafety, the three-way interaction effect of explanation modality, NDRT engagement, and technology suspicion was significant ($F = 5.09$, $p = 0.03$) as shown in Table 5. Therefore, H1b was supported. Specifically, we observed similar joint effects of these factors on anxiety as they had on perceived unsafety (Figure 5). Whether the individual had a high or low technology suspicion, there was no significant effect of modality on their perception of unsafety, as shown in slope 1 ($t = -0.13$; $p = 0.90$) and slope 3 ($t = -0.03$; $p = 0.98$), respectively. Individuals didn't perceive AVs as safer when given auditory or visual explanations while participating in the NDRT.

Nevertheless, without the NDRT, visual and auditory explanations were found to have different impacts on high- and low-suspicion individuals in terms of perceptions of unsafety. People with high technology suspicion, for example, perceived the AV as safer when they heard the explanation rather than read it (Slope 2; $t = 3.11$; $p = 0.01$). On the other hand, low-suspicion individuals preferred visual explanation, which resulted in lower feelings of unsafety as compared to auditory explanation (Slope 4; $t = -2.80$; $p = 0.01$).

C. SUMMARY OF THE FINDINGS

Together, the results suggest that explanation modality applications need to take into account individuals' technology suspicion and their engagement with an NDRT. Our findings clearly demonstrate that auditory explanations reduce anxiety and unsafety perception more effectively than visual explanations for high-suspicion individuals, especially in the absence of NDRT. Therefore, H1 was supported. The following section provides a more in-depth discussion of the

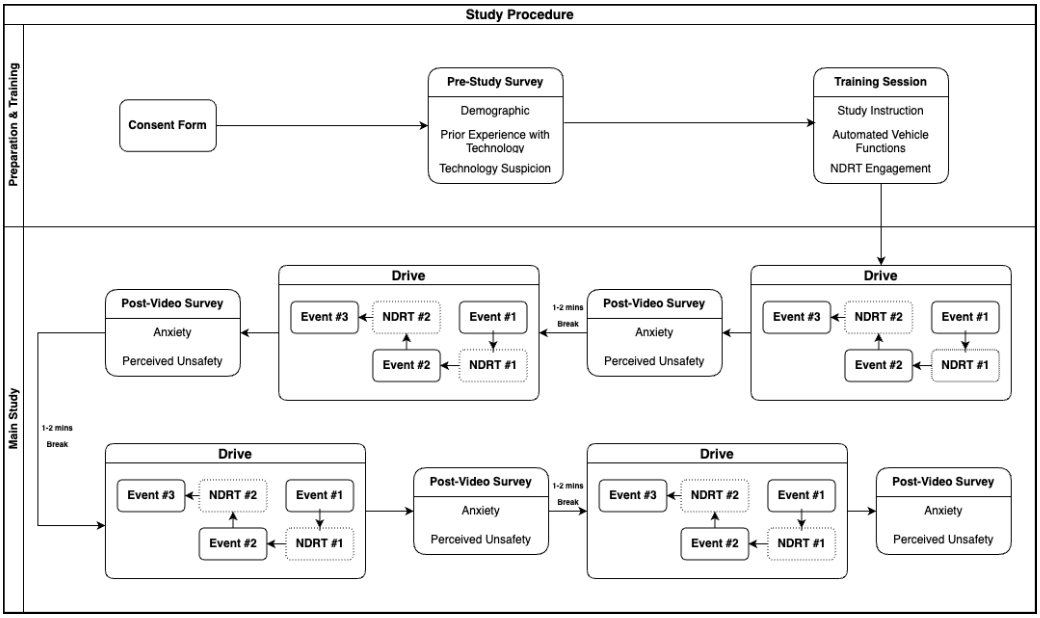

FIGURE 3. Study procedure.

TABLE 4. Descriptive Statistics.

Variable	Mean	Std.Dev	Mode	Median	Technology Suspicion	Anxiety	Unsafety
Technology Suspicion	2.76	0.91	3.20	2.60	1.00		
Anxiety	2.70	1.35	2.00	2.25	0.46	1.00	
Unsafety	2.16	0.78	1.00	2.20	0.33**	0.64**	1.00

**Correlation is significant at the 0.01 level (2-tailed).

TABLE 5. Interaction Effects of Explanation Modality, NDRT Engagement and Technology Suspicion with Anxiety and Unsafety.

	Anxiety			Unsafe		
	Coefficients	F	P	Coefficients	F	P
Intercept / Constant	2.64	125.26	<0.01	2.21	296.03	<0.01
Explanation Modality (EM)	0.18	0.88	0.35	0.01	0.01	0.91
Secondary Task Engagement (STE)	-0.04	0.47	0.83	-0.10	1.10	0.30
Technology Suspicion (TS)	-0.27	0.67	0.42	0.21	1.38	0.25
EM x STE	-0.04	0.03	0.87	-0.03	0.06	0.81
EM x TS	0.62	5.35	0.02*	0.41	9.52	<0.01**
STE x TS	0.49	3.33	0.07	0.10	0.52	0.48
EM x STE x TS	-0.80	4.38	0.04*	-0.42	5.09	0.03*

Notes: ** significance at 0.01 level; * significance at 0.05 level.

findings and their contributions to the literature, along with the study's limitations.

VI. DISCUSSION

A. AV EXPLANATION AND ACCEPTANCE

Previous research has shown that providing appropriate explanations can increase people's positive emotional response and acceptance of autonomous vehicles (AVs) [1]. Specifically, when the AV provides a why-only explanation for its actions before taking them, it can improve attitudes towards the technology, decrease negative emotions like anxiety, and enhance driving safety. This highlights the importance of effective explanations in promoting AV acceptance and the need to design explanations that are understandable and intuitive for users.

Our study builds upon existing research by demonstrating that the modality of explanation is a crucial factor in determining the effectiveness of AV explanations. Our findings suggest that the success of AV explanations is contingent upon both the task at hand and the delivery modality. Specifically, our results indicate that auditory explanations are

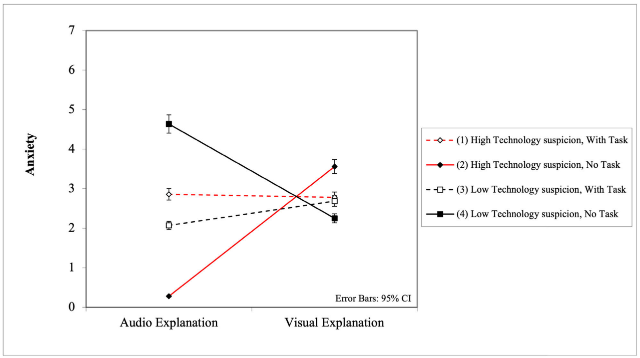

FIGURE 4. The three-way interaction of explanation modality, secondary task engagement, and technology suspicion with anxiety.

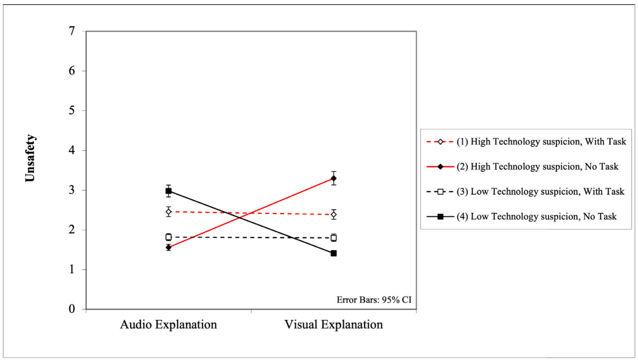

FIGURE 5. The three-way interaction of explanation modality, secondary task engagement, and technology suspicion with unsafety perception.

effective in reducing emotional arousal when the individual is visually focused on monitoring the driving environment. Conversely, visual explanations are effective in reducing anxiety and fear when the individual has less attention available for monitoring driving.

To conclude, our study highlights the crucial role of explanation modality in determining the effectiveness of AV explanations. Our results support the existing literature on the importance of explanations and emphasize the need to carefully consider the modality used in the design of AV explanations.

B. AUDITORY AV EXPLANATION, TECHNOLOGY SUSPICION, AND NDRT ENGAGEMENT

Studies have shown that auditory cues are effective in supporting attention orientation, facilitating time-shared driving, and promoting safe driving in levels 1-3 of driving automation. In fact, auditory information has been found to be superior to visual cues in these contexts [9], [10], [31], [35], [36], [37], [39]. However, it's worth noting that auditory explanations may cause heightened levels of stress and emotional arousal [11], [12], [13].

In addition, the effectiveness of the auditory modality is moderated by the presence of NDRTs, which have produced mixed results. While some studies have shown that auditory cues are particularly relevant to safe driving when an NDRT is present and elicit positive evaluations of the vehicle, such as satisfaction, usefulness, effectiveness, and safety [34], [40], [41], [42], [43], others suggest that auditory information in multitasking contexts does not have a significant effect on overall driving performance [67].

Our study sheds light on the influence of technology suspicion and NDRT engagement on the effectiveness of auditory explanations. Prior research has produced mixed results, and our findings explain some of these inconsistencies. First, our study confirms the benefits of auditory explanations in reducing anxiety and perceptions of unsafety, particularly for individuals who monitor their driving environment closely due to high technology suspicion [9], [10]. We found that auditory explanations are more effective for individuals with high technology suspicion, who have a greater need for effective explanations when they are not engaged in an NDRT. Auditory explanations allow individuals to monitor their driving environment while perceiving the message simultaneously, creating a driving experience with less anxiety and lower feelings of unsafety. Second, our research supports previous findings that auditory information can elicit stress and startle [11], [12], [34]. Specifically, individuals with low technological suspicion perceived auditory explanations negatively, resulting in higher anxiety and perceptions of unsafety. The use of auditory explanations can appear intrusive and ineffective to low-suspicion individuals, who are less likely to actively process the information, especially when no NDRT is present. Lastly, our findings suggest that the auditory modality may not be as effective in a multitasking environment [67]. Both high- and low-suspicion individuals experienced similar levels of anxiety and perceived unsafety when presented with auditory and visual explanations while engaged in an NDRT. This result could be attributed to the attention paid to the NDRT, which is likely to override the influence of the explanation modality on individuals' emotional responses. Therefore, the levels of anxiety and perceived unsafety were independent of the explanation modality.

C. VISUAL AV EXPLANATION, TECHNOLOGY SUSPICION, AND NDRT ENGAGEMENT

Previous research has shown that the visual modality is not as beneficial as auditory cues in supporting safe driving, but it does have some advantages. These include the ability to maintain continuous awareness of the surrounding environment [10], improve ease of use and transparency [44], and require short recognition times [17]. However, in a multitasking context, the benefits of the visual modality are limited. Researchers have found that visual alerts may not be sufficient to direct individuals' attention and improve driving performance in the presence of an NDRT [37], [38], [47].

By incorporating technology suspicion levels and NDRT engagement into prior research, our study extends and supports previous findings. The results suggest that visual explanations are preferred by individuals with low levels of technology suspicion, as they result in lower levels of anxiety and perceived unsafety. As these individuals may not actively monitor traffic and AV performance, visual explanations can support their situational awareness without being intrusive. However, for individuals with high levels of technology suspicion, visual explanations can increase anxiety and feelings of insecurity. The visual explanation competes for attentional resources, interfering with the primary task, based on the MRT and bottleneck theory of attention [39], [50], [52], [53], [54], [55]. Therefore, visual explanations are less useful for these individuals, especially when not engaged in an NDRT. Our study also confirms that the visual modality is of no importance in multitasking contexts [37], [38], [47]. In the presence of an NDRT, the impact of visual explanations on individuals is suppressed, similar to that of the auditory modality. As individuals primarily focus on the NDRT, they give little attention to the explanation, to the extent that it does not affect their emotions.

D. DESIGN IMPLICATIONS

The findings of this study have several implications for AV design. First, the modality of AV explanations should be considered, particularly when individuals are not engaged in an NDRT. Depending on the level of effort required to monitor the driving environment, individuals may receive explanations actively or passively. Auditory explanations act as a "push" technique for driving, allowing individuals to actively direct their attention to critical situations without interfering with their visual tasks. On the other hand, visual explanations are a passive, pull-like technique for acquiring information. In situations where individuals require AV action explanations, they should be able to focus their attention on the interface and receive the necessary information without disturbance.

Our study also offers specific guidelines for user interface (UI) designers of AVs to take into account when addressing users' diverse technology suspicion levels during the design process. By tailoring explanation modalities to users' technology suspicion levels, designers can effectively alleviate user anxiety and perceptions of unsafety. Our findings indicate that individuals with high levels of technology suspicion are more receptive to auditory explanations, such as voice-overs, whereas users with low levels of technology suspicion favor visual explanations like pop-up messages or icons. As a result, the initial step for designers should be to actively assess users' technology suspicion levels in order to effectively apply these findings.

Existing approaches for evaluating individual technology suspicion levels involve self-report measures found in prior literature, such as the Complacency Potential Rating Scale [26], Suspicion Propensity Index [68], and Privacy Concerns Scale [69], [70]. These measures typically require participants to rate their distrust or skepticism toward new technology. However, this method has several drawbacks, including misinterpretation of survey questions, biased answers, and survey fatigue [71], [72]. An alternative approach entails assessing users' technology suspicion levels by observing their behavior in real-time before or during their interactions with the AV. Previous research has identified patterns in the behaviors of individuals with high levels of technology suspicion, such as avoiding or resisting new technology [73], [74], expressing distrust or skepticism [75], [76], or harboring concerns about privacy and security [77], [78]. By identifying these patterns, it may be feasible to develop models that predict or estimate technology suspicion levels for individuals or groups.

An additional strategy involves creating pre-built profiles of technology suspicion based on user demographics. This method can offer a starting point for designers, depending on the predictability linked to each demographic category. For instance, Zhang et al. underscore the significance of individual differences, including demographics, in influencing attitudes and behaviors related to AVs [79]. However, this approach may result in inaccurate estimates for individual users who deviate from the pre-built profiles. A more comprehensive method would combine both pre-built profiles and real-time assessments. By integrating multiple sources of information, designers can develop a more precise and adaptable system for addressing users' technology suspicion levels, ultimately improving the user experience and promoting trust in AV technology.

In conclusion, providing explanations alone is not sufficient for individuals to feel comfortable with AV technology. The explanation modality is equally important and must be intuitive for users. AV designers must consider individuals' technology suspicion levels and NDRT engagement when selecting the appropriate explanation modality. It is imperative to provide explanations in a way that is easily understandable and not intrusive to support safe and effective AV use.

E. RESEARCH CONTRIBUTIONS

The present study makes a significant threefold contribution to the understanding of AV explanations and their effectiveness. First, this research emphasizes the essential role that explanation modality plays in determining the success of AV explanations. It demonstrates that the effectiveness of these explanations relies on the task being executed and the chosen delivery modality. In particular, the study's findings align with the Multiple Resource Theory (MRT), suggesting that auditory explanations effectively diminish emotional arousal when an individual is occupied with monitoring the driving environment. In contrast, visual explanations efficiently alleviate anxiety and fear when the individual's attention is less focused on driving.

Second, the study illuminates the impact of technology suspicion and NDRT engagement on the efficacy of AV explanations. When individuals are not involved in NDRT, their preference for explanation modalities often varies based on their level of technology suspicion. Individuals who exhibit high levels of technology suspicion typically lean towards auditory explanations, while those with low technology suspicion levels show a preference for visual explanations.

However, the dynamics change when an NDRT is introduced. In this scenario, the research supports the bottleneck theory of attention, which suggests that our capacity to process information from various sources at the same time has its limits. Consequently, in multitasking situations involving an NDRT, the influence of explanation modality becomes insignificant, regardless of an individual's technology suspicion levels.

The third contribution of this study is the provision of concrete guidelines for AV designers to develop AVs that cater to users with varying levels of technology suspicion. By proactively monitoring users' technology suspicion and offering clear, appropriate explanation modalities, designers can create AVs that successfully mitigate mistrust and negative perceptions towards AVs, enhance user experience, and promote technology adoption.

F. LIMITATIONS AND FUTURE RESEARCH

There are several limitations to this study that should be considered. First, the participants in this study were recruited from a university-related subject pool, which may not be representative of the general population's knowledge and experience in the AV field. Second, while the experimental setting had high internal validity, there are external validity limitations, and future studies should be conducted in field settings to improve generalizability. Third, participants may have engaged in hypothesis guessing and altered their reactions and responses based on what they thought the researcher desired, although we found no evidence of this in our study. Fourth, this study focuses on AVs, specifically those with level 4 and above automation systems, which often require more detailed explanations than simpler alternatives. Nevertheless, our findings could potentially be relevant in the context of Advanced Driver Assistance Systems (ADAS). Future research should explore this area to enhance the understanding and design of more sophisticated driving systems. Furthermore, although human-machine interfaces in prior studies deliver alerts through multi-modal feedback including visual, auditory, and haptic cues, this study only addresses unimodal communication (visual vs. auditory). As the content of AV explanations in this study is longer and more complex than that of alerts, we exclude tactile feedback due to its inadequacy to deliver complex information [80], [81]. Moreover, because prior research in AV explanation utilized only uni-modal to explain, and one of the motivations of this study was to reveal and explain the mixed results of prior literature and examine the potential impact of explanation modality, we have chosen only visual and auditory modalities as uni-modal explanation modes. Future studies should focus on human-machine interfaces using multi-modal, haptic feedback, which may be preferred for the design of future automated vehicles. Furthermore, this study did not analyze other features associated with AV explanations, such as the definition, generation, selection, and evaluation of alternative courses of action for the individual. Future studies should examine these and other possible attributes related to AV explanations. In summary, more research is needed to fully understand the impact of AV explanations and how to design them effectively.

VII. CONCLUSION

In this study, we investigated the impact of modality, NDRT engagement, and individuals' technology suspicion on the effectiveness of AV explanations, specifically their ability to reduce anxiety and perceptions of unsafety. To our knowledge, this is the first study to examine the effect of modality on explanation effectiveness in Level 5 automation. Our results highlight the importance of considering the interaction between modality and individuals' states and traits in determining the effectiveness of AV explanations. We also identified three properties of explanation (content, time, and modality) that influence its effectiveness, and future research could further explore how communication style affects explanation effectiveness. Overall, our study contributes to a better understanding of the factors that affect AV explanation effectiveness, which is essential for the design of future automated vehicles.

REFERENCES

[1] Q. Zhang, X. J. Yang, and L. P. Robert, "What and when to explain? A survey of the impact of explanation on attitudes toward adopting automated vehicles," *IEEE Access*, vol. 9, pp. 159533–159540, 2021.

[2] SAE International, "Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles," *SAE Int.*, vol. 4970, no. 724, pp. 1–5, 2018.

[3] N. Du, J. Haspiel, Q. Zhang, D. Tilbury, A. K. Pradhan, X. J. Yang, and L. P. Robert, "Look who's talking now: Implications of AV's explanations on driver's trust, AV preference, anxiety and mental workload," *Transp. Res. C, Emerg. Technol.*, vol. 104, pp. 428–442, Jul. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X18313640

[4] Y. Forster, F. Naujoks, and A. Neukum, "Increasing anthropomorphism and trust in automated driving functions by adding speech output," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2017, pp. 365–372.

[5] J. Koo, D. Shin, M. Steinert, and L. Leifer, "Understanding driver responses to voice alerts of autonomous car operations," *Int. J. Veh. Des.*, vol. 70, no. 4, pp. 377–392, 2016, doi: 10.1504/IJVD.2016.076740.

[6] J. Koo, J. Kwac, W. Ju, M. Steinert, L. Leifer, and C. Nass, "Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance," *Int. J. Interact. Design Manuf.*, vol. 9, no. 4, pp. 269–275, Nov. 2015, doi: 10.1007/s12008-014-0227-2.

[7] M. Körber, L. Prasch, and K. Bengler, "Why do I have to drive now? Post hoc explanations of takeover requests," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 60, no. 3, pp. 305–323, May 2018, doi: 10.1177/0018720817747730.

[8] S. Thill, P. E. Hemeren, and M. Nilsson, "The apparent intelligence of a system as a factor in situation awareness," in *Proc. IEEE Int. Inter-Disciplinary Conf. Cognit. Methods Situation Awareness Decis. Support (CogSIMA)*, Mar. 2014, pp. 52–58.

[9] Y. Cao, A. Mahr, S. Castronovo, M. Theune, C. Stahl, and C. A. Müller, "Local danger warnings for drivers: The effect of modality and level of assistance on driver reaction," in *Proc. 15th Int. Conf. Intell. User Interface*, Feb. 2010, pp. 239–248.

[10] D. J. Wheatley and J. B. Hurwitz, "The use of a multi-modal interface to integrate in-vehicle information presentation," in *Proc. 1st Int. Driving Symp. Hum. Factors Driver Assessment, Training Vehicle Design, Driving Assessment*, 2001, pp. 1–12.

[11] T. A. Dingus, D. V. McGehee, N. Manakkal, S. K. Jahns, C. Carney, and J. M. Hankey, "Human factors field evaluation of automotive headway maintenance/collision warning devices," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 39, no. 2, pp. 216–229, Jun. 1997.

[12] N. Gupta, A. M. Bisantz, and T. Singh, "The effects of adverse condition warning system characteristics on driver performance: An investigation of alarm signal type and threshold level," *Behav. Inf. Technol.*, vol. 21, no. 4, pp. 235–248, Jan. 2002.

[13] M. A. Nees, B. Helbein, and A. Porter, "Speech auditory alerts promote memory for alerted events in a video-simulated self-driving car ride," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 58, no. 3, pp. 416–426, May 2016.

[14] A. H. Jamson, N. Merat, O. M. J. Carsten, and F. C. H. Lai, "Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions," *Transp. Res. Part C: Emerg. Technol.*, vol. 30, pp. 116–125, May 2013.

[15] S. Carley, R. M. Krause, B. W. Lane, and J. D. Graham, *Transp. Res. D, Transp. Environ.*, vol. 18, p. 39, Jan. 2013.

[16] K. Zeeb, A. Buchner, and M. Schrauf, "Is take-over time all that matters? The impact of visual-cognitive load on driver take-over quality after conditionally automated driving," *Accident Anal. Prevention*, vol. 92, pp. 230–239, Jul. 2016.

[17] I. Politis, S. Brewster, and F. Pollick, "To beep or not to beep? Comparing abstract versus language-based multimodal driver displays," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, Apr. 2015, pp. 3971–3980, doi: 10.1145/2702123.2702167.

[18] Y.-C. Liu, "Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems," *Ergonomics*, vol. 44, no. 4, pp. 425–442, Mar. 2001, doi: 10.1080/00140130010011369.

[19] P. Bobko, A. J. Barelka, and L. M. Hirshfield, "The construct of state-level suspicion: A model and research agenda for automated and information technology (IT) contexts," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 56, no. 3, pp. 489–508, May 2014, doi: 10.1177/0018720813497052.

[20] B. M. Muir, "Trust between humans and machines, and the design of decision aids," *Int. J. Man-Mach. Stud.*, vol. 27, nos. 5–6, pp. 527–539, Nov. 1987. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020737387800135

[21] R. Parasuraman, R. Molloy, and I. L. Singh, "Performance consequences of automation-induced 'complacency,'" *Int. J. Aviation Psychol.*, vol. 3, no. 1, pp. 1–23, Jan. 1993, doi: 10.1207/s15327108ijap0301_1.

[22] J. Benamati and M. A. Serva, "Trust and distrust in online banking: Their role in developing countries," *Inf. Technol. Develop.*, vol. 13, no. 2, pp. 161–175, Apr. 2007, doi: 10.1002/itdj.20059.

[23] J. Lu, C.-S. Yu, and C. Liu, "Facilitating conditions, wireless trust and adoption intention," *J. Comput. Inf. Syst.*, vol. 46, no. 1, pp. 17–24, Sep. 2005.

[24] J. B. Lyons, C. K. Stokes, K. J. Eschleman, G. M. Alarcon, and A. J. Barelka, "Trustworthiness and IT suspicion: An evaluation of the nomological network," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 53, no. 3, pp. 219–229, Jun. 2011.

[25] E. Rovira, K. McGarry, and R. Parasuraman, "Effects of imperfect automation on decision making in a simulated command and control task," *Hum. Factors, J. Human Factors Ergonom. Soc.*, vol. 49, no. 1, pp. 76–87, Feb. 2007.

[26] I. L. Singh, R. Molloy, and R. Parasuraman, "Automation-induced 'complacency': Development of the complacency-potential rating scale," *Int. J. Aviation Psychol.*, vol. 3, no. 2, pp. 111–122, Apr. 1993, doi: 10.1207/s15327108ijap0302_2.

[27] S. O.-R. A. V. S. Committee. (2021). *J3016C: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles—SAE International*. [Online]. Available: https://www.sae.org/standards/content/j3016_202104/

[28] P. Ruijten, J. Terken, and S. Chandramouli, "Enhancing trust in autonomous vehicles through intelligent user interfaces that mimic human behavior," *Multimodal Technol. Interact.*, vol. 2, no. 4, p. 62, Sep. 2018.

[29] Q. Zhang, X. J. Yang, and L. P. Robert, "Drivers' age and automated vehicle explanations," *Sustainability*, vol. 13, no. 4, p. 1948, Feb. 2021.

[30] G. Wiegand, M. Schmidmaier, T. Weber, Y. Liu, and H. Hussmann, "I drive–you trust: Explaining driving behavior of autonomous cars," in *Proc. Extended Abstr. CHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, May 2019, pp. 1–6, doi: 10.1145/3290607.3312817.

[31] F. Naujoks, C. Mai, and A. Neukum, "The effect of urgency of take-over requests during highly automated driving under distraction conditions," *Adv. Hum. Aspects Transp.*, vol. 7, p. 431, Jul. 2014.

[32] F. Karray, M. Alemzadeh, J. A. Saleh, and M. N. Arab, "Human–computer interaction: Overview on state of the art," *Int. J. Smart Sens. Intell. Syst.*, vol. 1, no. 1, pp. 137–159, Jan. 2008.

[33] K. Mahadevan, S. Somanath, and E. Sharlin, "Communicating awareness and intent in autonomous vehicle-pedestrian interaction," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, Apr. 2018, pp. 1–12.

[34] M. A. Nees and B. N. Walker, "Auditory displays for in-vehicle technologies," *Rev. Hum. Factors Ergonom.*, vol. 7, no. 1, pp. 58–99, Sep. 2011.

[35] N. O. Bernsen and L. Dybkjaer, "Exploring natural interaction in the car," in *Proc. CLASS Workshop Natural Interactivity Intell. Interact. Inf. Represent.*, vol. 2, 2001, pp. 1–12.

[36] B. Donmez, L. N. Boyle, J. D. Lee, and D. V. McGehee, "Drivers' attitudes toward imperfect distraction mitigation strategies," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 9, no. 6, pp. 387–398, Nov. 2006.

[37] S. Petermeijer, F. Doubek, and J. de Winter, "Driver response times to auditory, visual, and tactile take-over requests: A simulator study with 101 participants," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 1505–1510.

[38] S. H. Yoon, Y. W. Kim, and Y. G. Ji, "The effects of takeover request modalities on highly automated car control transitions," *Accident Anal. Prevention*, vol. 123, pp. 150–158, Feb. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001457518310285

[39] C. D. Wickens, "Multiple resources and mental workload," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 50, no. 3, pp. 449–455, Jun. 2008.

[40] J.-H. Lee and C. Spence, "Assessing the benefits of multimodal feedback on dual-task performance under demanding conditions," *People Comput. Culture, Creativity, Interact.*, vol. 22, pp. 185–192, Jan. 2008.

[41] G. R. Widmann, J. Salinger, C. A. Green, F. Bolourchi, M. Nanjundaiah, R. Prieto, R. E. Llaneras, T. Yang, T. Kaufmann, and Y. Chen, "A flexible system for studying driver visual attentiveness while using semi-autonomous driving systems," in *Proc. 20th ITS World CongressITS Japan*, 2013, pp. 1–15.

[42] N. Merat, A. H. Jamson, F. C. H. Lai, M. Daly, and O. M. J. Carsten, "Transition to manual: Driver behaviour when resuming control from a highly automated vehicle," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 27, pp. 274–282, Nov. 2014.

[43] C. Lee, C. Ward, M. Raue, L. D'Ambrosio, and J. F. Coughlin, "Age differences in acceptance of self-driving cars: A survey of perceptions and attitudes," in *Human Aspects of IT for the Aged Population. Aging, Design and User Experience*, J. Zhou and G. Salvendy, Eds. Cham, Switzerland: Springer, 2017, pp. 3–13.

[44] N. Du, F. Zhou, D. Tilbury, L. P. Robert, and X. J. Yang, "Designing alert systems in takeover transitions: The effects of display information and modality," in *Proc. 13th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, Sep. 2021, pp. 173–180, doi: 10.1145/3409118.3475155.

[45] L. Avetisyan, J. Ayoub, and F. Zhou, "Investigating explanations in conditional and highly automated driving: The effects of situation awareness and modality," 2022, *arXiv:2207.07496*.

[46] Y. Cao, S. Castronovo, A. Mahr, and C. Müller, "On timing and modality choice with local danger warnings for drivers," in *Proc. 1st Int. Conf. Automot. User Interface Interact. Veh. Appl.*, Sep. 2009, pp. 75–78.

[47] J. Radlmayr, C. Gold, L. Lorenz, M. Farid, and K. Bengler, "How traffic situations and non-driving related tasks affect the take-over quality in highly automated driving," *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, vol. 58, no. 1, pp. 2063–2067, Sep. 2014.

[48] G. Huang and B. J. Pitts, "Takeover requests for automated driving: The effects of signal direction, lead time, and modality on takeover performance," *Accident Anal. Prevention*, vol. 165, Feb. 2022, Art. no. 106534.

[49] C. D. Wickens, *Processing Resources and Attention*. Boca Raton, FL, USA: CRC Press, 2020.

[50] B. Wahn and P. König, "Is attentional resource allocation across sensory modalities task-dependent?" *Adv. Cognit. Psychol.*, vol. 13, no. 1, pp. 83–96, Mar. 2017.

[51] N. B. Sarter, "Multiple-resource theory as a basis for multimodal interface design: Success stories, qualifications, and research needs," in *Attention: From Theory to Practice*. New York, NY, USA: Oxford Univ. Press, 2007, pp. 187–195.

[52] R. E. Smith and L. M. Buchholz, "Multiple resource theory and consumer processing of broadcast advertisements: An involvement perspective," *J. Advertising*, vol. 20, no. 3, pp. 1–7, Sep. 1991.
[53] H. Pashler, "Dual-task interference in simple tasks: Data and theory," *Psychol. Bull.*, vol. 116, no. 2, pp. 220–244, 1994.
[54] C. D. Wickens and J. S. McCarley, *Applied Attention Theory*. Boca Raton, FL, USA: CRC Press, 2019.
[55] D. E. Broadbent, *Perception and Communication*. Amsterdam, The Netherlands: Elsevier, 2013.
[56] S. Cohen, "Perceived stress in a probability sample of the United States," in *The Social Psychology of Health* (The Claremont Symposium on Applied Social Psychology). Thousand Oaks, CA, USA: Sage, 1988, pp. 31–67.
[57] C. Nass, I.-M. Jonsson, H. Harris, B. Reaves, J. Endo, S. Brave, and L. Takayama, "Improving automotive safety by pairing driver emotion and car voice emotion," in *Proc. CHI Extended Abstr. Hum. Factors Comput. Syst.*, Apr. 2005, pp. 1973–1976.
[58] P. Koopman and M. Wagner, "Toward a framework for highly automated vehicle safety validation," SAE Tech. Paper 0148-7191, 2018.
[59] B. E. Hayes, J. Perander, T. Smecko, and J. Trask, "WITHDRAWN: Reprint of 'measuring perceptions of workplace safety: Development and validation of the work safety scale,"' *J. Saf. Res.*, vol. 5, pp. 145–161, Jul. 2013.
[60] M. G. Lenné, T. J. Triggs, C. M. Mulvihill, M. A. Regan, and B. F. Corben, "Detection of emergency vehicles: Driver responses to advance warning in a driving simulator," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 50, no. 1, pp. 135–144, Feb. 2008.
[61] N. Merat and A. H. Jamson, "How do drivers behave in a highly automated car?" in *Proc. Driving Assessment Conf.*, vol. 5. Iowa City, IA, USA: Univ. of Iowa, Public Policy Center, 2009, pp. 514–521.
[62] D. Miller and W. Ju, "Does the first officer concur? Shared control with smart vehicle systems," in *Proc. Adjunct Proc. 6th Int. Conf. Automot. User Interface Interact. Veh. Appl.*, Sep. 2014, pp. 1–6.
[63] L. J. Molnar, L. H. Ryan, A. K. Pradhan, D. W. Eby, R. M. S. Louis, and J. S. Zakrajsek, "Understanding trust and acceptance of automated vehicles: An exploratory simulator study of transfer of control between automated and manual driving," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 58, pp. 319–328, Oct. 2018.
[64] T. Rezvani, K. Driggs-Campbell, D. Sadigh, S. S. Sastry, S. A. Seshia, and R. Bajcsy, "Towards trustworthy automation: User interfaces that convey internal and external awareness," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst.*, Nov. 2016, pp. 682–688.
[65] G. W. Wille, "A stepwise procedure for the empirical assessment of latent variable models," Universiteit van Port Elizabeth, 1996.
[66] J. Raubenheimer, "An item selection procedure to maximise scale reliability and validity," *SA J. Ind. Psychol.*, vol. 30, no. 4, pp. 59–64, Oct. 2004.
[67] R. Mohebbi, R. Gray, and H. Z. Tan, "Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 51, no. 1, pp. 102–110, Feb. 2009, doi: 10.1177/0018720809333517.
[68] B. Levi and G. Loeben, "Index of suspicion: Feeling not believing," *Theor. Med. Bioethics*, vol. 25, no. 4, pp. 277–310, 2004.
[69] S. Egelman and E. Peer, "The myth of the average user: Improving privacy and security systems through individualization," in *Proc. New Secur. Paradigms Workshop*, Sep. 2015, pp. 16–28.
[70] S. Stieger, C. Burger, M. Bohn, and M. Voracek, "Who commits virtual identity suicide? Differences in privacy concerns, internet addiction, and personality between Facebook users and quitters," *Cyberpsychol., Behav., Social Netw.*, vol. 16, no. 9, pp. 629–634, Sep. 2013.
[71] J. A. Krosnick, "Survey research," *Annu. Rev. Psychol.*, vol. 50, no. 1, pp. 537–567, 1999.
[72] B. C. Choi and A. W. Pak, "Peer reviewed: A catalog of biases in questionnaires," in *Preventing Chronic Disease*, vol. 2, no. 1. Atlanta, GA, USA: Centers for Disease Control and Prevention, 2005.
[73] L. Lapointe and S. Rivard, "A multilevel model of resistance to information technology implementation," *MIS Quart.*, vol. 10, pp. 461–491, Sep. 2005.
[74] M. W. Bauer, *Resistance to New Technology: Nuclear Power, Information Technology and Biotechnology*. Cambridge, U.K.: Cambridge Univ. Press, 1995.
[75] K. Stenner, E. R. Frederiks, E. V. Hobman, and S. Cook, "Willingness to participate in direct load control: The role of consumer distrust," *Appl. Energy*, vol. 189, pp. 76–88, Jan. 2017.
[76] Y. Tsfati and J. N. Cappella, "Do people watch what they do not trust? Exploring the association between news media skepticism and exposure," *Commun. Res.*, vol. 30, no. 5, pp. 504–529, Oct. 2003.
[77] S. Kokolakis, "Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon," *Comput. Secur.*, vol. 64, pp. 122–134, Jan. 2017.
[78] S. Pearson, "Taking account of privacy when designing cloud computing services," in *Proc. ICSE Workshop Softw. Eng. Challenges Cloud Comput.*, Mar. 2009, pp. 44–52.
[79] Q. Zhang, X. J. Yang, and L. P. Robert, "Individual differences and expectations of automated vehicles," *Int. J. Hum.-Comput. Interact.*, vol. 2021, pp. 1–12, Sep. 2021, doi: 10.1080/10447318.2021.1970431.
[80] M. S. Prewett, L. Yang, F. R. B. Stilson, A. A. Gray, M. D. Coovert, J. Burke, E. Redden, and L. R. Elliot, "The benefits of multimodal information: A meta-analysis comparing visual and visual-tactile feedback," in *Proc. 8th Int. Conf. Multimodal Interface*, Nov. 2006, pp. 333–338.
[81] J. Rantala, R. Raisamo, J. Lylykangas, V. Surakka, J. Raisamo, K. Salminen, T. Pakkanen, and A. Hippula, "Methods for presenting Braille characters on a mobile device with a touchscreen and tactile feedback," *IEEE Trans. Haptics*, vol. 2, no. 1, pp. 28–39, Jan. 2009.

QIAONING ZHANG received the master's degree in industrial and operations engineering from the University of Michigan, where she is currently pursuing the Ph.D. degree with the School of Information. She is a member of the Michigan Autonomous Vehicle Research Intergroup Collaboration (MAVRIC).

CONNOR ESTERWOOD (Graduate Student Member, IEEE) is currently pursuing the Ph.D. degree with the School of Information, University of Michigan. His recent work has focused on trust repair, with an emphasis on apologies, explanations, denials, and promises. His current research interests include human–robot interaction and human-AI teams.

ANUJ K. PRADHAN received the B.E. degree in mechanical engineering from India and the M.S. and Ph.D. degrees in industrial engineering from the University of Massachusetts Amherst. He was a Visiting Fellow with the National Institutes of Health and an Assistant Research Scientist with the University of Michigan Transportation Research Institute. He is currently an Assistant Professor with the Mechanical and Industrial Engineering Department, University of Massachusetts Amherst. He has authored more than 50 articles. He is a member of HFES, AAAM, ACM, and various TRB committees.

DAWN TILBURY (Fellow, IEEE) is currently the inaugural Department Chair of robotics and the Herrick Professor of engineering with the University of Michigan, Ann Arbor. She has published more than 200 papers in refereed journals and conference proceedings. She is a fellow of ASME and a Life Member of SWE.

X. JESSIE YANG received the Ph.D. degree in mechanical and aerospace engineering (human factors) from Nanyang Technological University, Singapore, in 2014. She is currently an Associate Professor with the Department of Industrial and Operations Engineering and the School of Information (by courtesy), University of Michigan, Ann Arbor.

LIONEL P. ROBERT JR. (Senior Member, IEEE) is currently a Professor with the School of Information and a Core Faculty Member of the Robotics Institute, University of Michigan. He is also the Director of the Michigan Autonomous Vehicle Research Intergroup Collaboration (MAVRIC) and an Association for Information Systems Distinguished Member Cum Laude.